

Федеральное государственное автономное  
образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
Институт космических и информационных технологий  
Кафедра «Вычислительная техника»

УТВЕРЖДАЮ

Заведующий кафедрой

А.И. Легалов

\_\_\_\_\_  
подпись

«\_\_\_\_\_» \_\_\_\_\_ 2016 г.

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

**Генерация учебных заданий на основе текста с использованием  
порождающих грамматик над многомерными данными**

09.04.01 Информатика и вычислительная техника

09.04.01.04 Технология разработки программного обеспечения

Научный руководитель

к.т.н.

Д.В. Личаргин

\_\_\_\_\_  
подпись, дата

Выпускник

А.А. Усова

\_\_\_\_\_  
подпись, дата

Нормоконтроллер

к.т.н., доцент

В.И. Иванов

\_\_\_\_\_  
подпись, дата

Рецензент

д.ф.-м.н.,

К.В. Сафонов

доцент

\_\_\_\_\_  
подпись, дата

Красноярск 2016

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	4
1 Обзор аналогичных разработок .....	6
1.1 Сравнение функционала аналогичных разработок .....	6
1.2 Онлайн программа Textivate .....	6
1.3 Программа QuickWorksSheets.....	9
1.3 Программа WorksSheet Generator .....	12
1.4 Программа LearnClick.....	16
2 Методы и средства разработки .....	19
2.1 Порождающая грамматика над многомерными базами данных.....	19
2.2 OLAP технологии и применение многомерных баз данных при генерации осмысленной речи. ....	23
2.3 Использование метрики Хэмминга для оценки близости генерируемых заданий к тексту .....	25
2.4 Семантический шум.....	27
2.5 Математическая модель .....	29
3 Описание используемых программ в среде mouse auto clicker .....	33
3.1 Описание программы «Electronic Dictionary» .....	33
3.2 Описание программы «Acapela» .....	38
3.3 О системе электронного обучения Сибирского федерального университета.....	40
3.4 Описание работы программы «Mouse auto clicker» .....	43
4 Программное обеспечение естественно-языковых интерфейсов .....	46
4.1 Разработка системы «Генератор учебных заданий» и «Генератор порождаемых шаблонов» .....	46
4.2 Назначение системы «Генератор учебных заданий» .....	47
4.3 Функции системы «Генератор учебных заданий».....	48
4.4 Состав и структура системы «Генератор учебных заданий» .....	49
4.5 Входные и выходные данные .....	51
4.6 Функционирование системы «Генератор учебных заданий».....	52

4.7 Программная система «Генератор порождаемых шаблонов».....	53
4.7.1 Назначение программы «Генератор порождаемых шаблонов» .....	54
4.7.2 Функции программы «Генератор порождаемых шаблонов» .....	54
4.7.3 Структура программы «Генератор порождаемых шаблонов» .....	55
4.7.4 Входные и выходные данные программы «Генератор порождаемых шаблонов» .....	56
4.7.5 Описание работы программы «Генератор порождаемых шаблонов» .....	56
5 Описание разработанных программ.....	58
5.1 Визуальное отображение программы «Генератор порождаемых шаблонов» .....	58
5.2 Визуальное отображение программы «Генератор учебных заданий». ....	59
ЗАКЛЮЧЕНИЕ .....	64
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ .....	65
ПРИЛОЖЕНИЕ А .....	71
ПРИЛОЖЕНИЕ Б.....	74

## ВВЕДЕНИЕ

Задачей данного проекта является решение проблемы повышения эффективности языковой подготовки в техническом вузе, разработка средств генерации большего по мощности и более осмысленного множества учебных материалов в целях индивидуализации траекторий обучения.

Актуальность работы состоит в повышении эффективности работы преподавателей, снижении временных затрат, увеличении объемов генерируемых материалов, что обеспечит возможность построения множества индивидуальных траекторий обучения с целью повышения качества обучения. Разработка электронных обучающих ресурсов требует от преподавателя увеличения затрат временных ресурсов на создание учебного материала, что затрудняет реализацию индивидуальной траектории обучения студентов. В процессе создания современных систем электронного обучения требуется интеграция знаний из разнородных источников и решение задач тиражирования знаний.

Поэтому актуальна задача разработки ресурса, который может помочь преподавателям иностранного языка, генерировать тестовые материалы, для проверки знаний в целях повышения эффективности процесса составления учебных материалов.

**Цель работы:** Разработка программы повышения вариативности генерируемых фраз естественного языка, на основе гибридизации порождающих грамматик и многомерных баз данных. Обеспечение интеграции этой программы с сопутствующим комплексом программных продуктов на основе сценариев автоматической работы управляющих агентов.

**Научная новизна:** достижение определенного высокого уровня на основе использования модели гибридизации многомерных баз данных и порождающих грамматик.

**Задачи** диссертационной работы состоят в:

- разработке правил порождающей грамматики;

- разработке многомерных баз данных на основе бесплатного программного обеспечения;
- разработке модулей, позволяющих осуществить гибридизацию многомерных баз данных и порождающих грамматик;
- создании программного продукта позволяющего вариативно использовать данный унифицированный алгоритм;
- апробации данного программного продукта для разработки электронного курса «английский язык для академических целей» курса магистратуры.

**Назначение программы** – помочь преподавателям автоматизировать работу по разработке электронных образовательных курсов, а так же ускорить процесс разработки тестов на основе печатных материалов.

**Объектом исследования** являются автоматизированные обучающие системы иностранным языкам (на примере английского языка).

**Предмет исследований** – виды и методы автоматического анализа и автоматической компрессии текстов.

**Значение для теории.** Разработаны алгоритмы гибридизации порождающих грамматик и многомерных баз данных, применительно к генерации естественного языка и модели многомерного лексико-грамматического пространства.

**Практическая ценность.** Разработанная в диссертации программа генерации учебных заданий позволяет экономить временные и человеческие ресурсы при разработке учебных тестов.

**Область потенциального применения.** Разработанная программа может быть использована в Сибирском федеральном университете на кафедре «Разговорного иностранного языка» и в Сибирском государственном аэрокосмическом университете на кафедре «Прикладной математики» для автоматизации работы преподавателей при составлении учебных заданий к тексту.

## **1 Обзор аналогичных разработок**

### **1.1 Сравнение функционала аналогичных разработок**

Были рассмотрены как бесплатные, так и платные аналоги генераторов заданий к тексту. Некоторые генераторы расположены в сети интернет и позволяют преподавателям генерировать задания без установки программы, не выходя из браузера. Пользоваться такими генераторами намного проще, так как не нужно разбираться с установкой и настройкой.

### **1.2 Онлайн программа Textivate**

Textivate позволяет сгенерировать различные упражнения на основе введенного текста, но текст должен содержать не более 500 знаков. Так же есть возможность вставить ссылку на видео, аудио файл или изображение.

Главное окно изображено на рисунке 1. Здесь видно белое поле для ввода текста, а также ниже есть кнопки для сохранения заданий локально, для скачивания заданий, выбор шаблона, очистить окно от текста, поиск. На главной странице сайта предлагается ввести текст, не более 500 символов в специальное окно, а затем нужно нажать кнопку `textivate now` и далее откроется страница для выбора шаблона, которая изображена на рисунке 2.



Рисунок 1 – Главное окно программы textivate

Сайт предлагает различные варианты упражнений для работы с текстом: разделение пробелами слов, вписывание пропущенных букв и слов, расположение в правильном порядке отрывков, предложений, абзацев, анаграммы. На рисунке 2 показано окно, где выбирается шаблон для генерации заданий. Большое разнообразие заданий помогает индивидуализировать варианты, а так же выбрать уровень сложности.

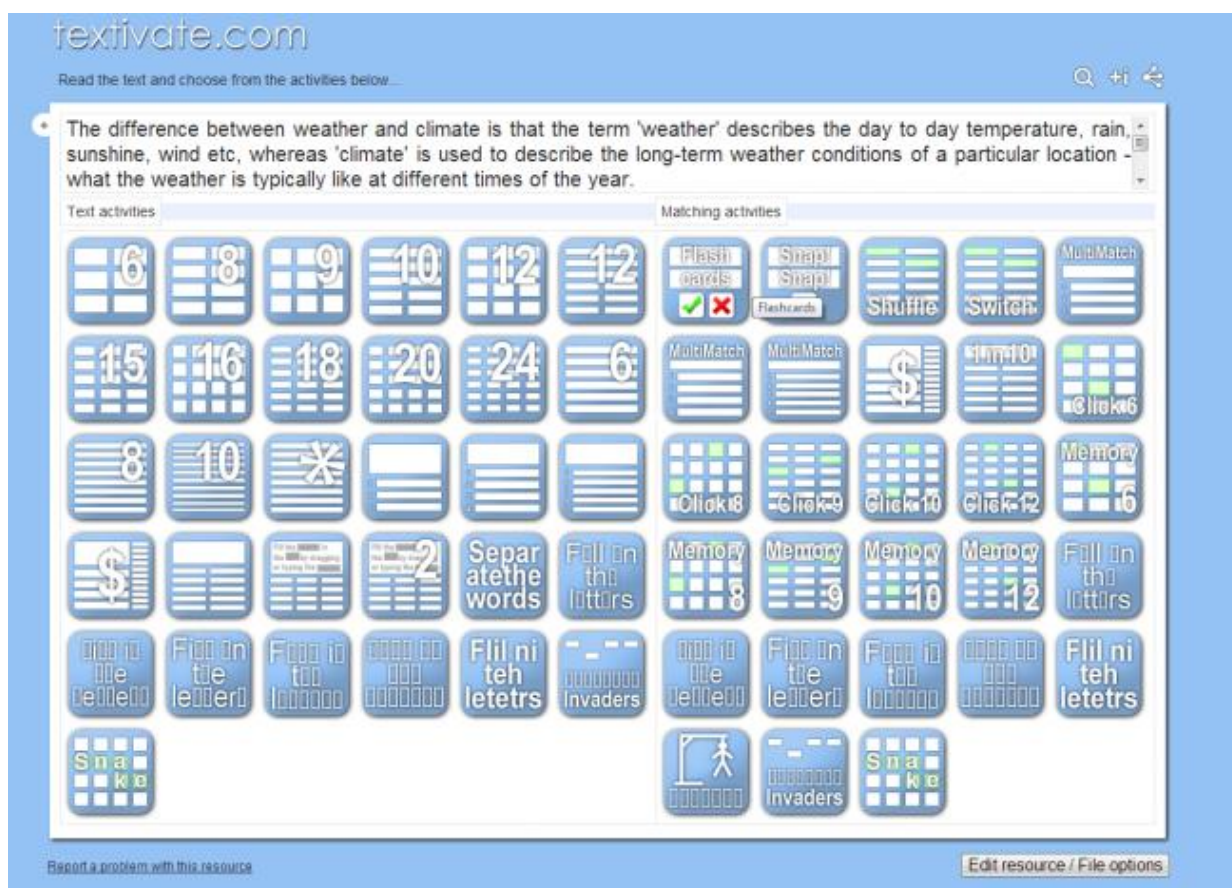


Рисунок 2 – Окно для выбора шаблона заданий

Помимо всего есть возможность в окне ввода текста создать вопросы на понимание текста, а так же слов с помощью тега `##match##`, разделив знаками `==` ответы и вопросы. Так же программа вставляет их автоматически, если список слов вставляется из Excel или Word.

Многие функции, например, сохранение материалов на диск, доступны только при платной подписке.

Textivate вполне может стать хорошим помощником преподавателя при подготовке к занятиям.



### 1.3 Программа QuickWorksheets

Данная программа позволяет создать упражнения, игры и задания для учащихся. Сайт предлагает 22 варианта рабочих распечаток, 3 из которых бесплатные, а остальные доступны в pro-версии продукта.

На основе нужных материалов можно создать кроссворды, игры с поиском слов в поле букв, игры на соотнесение, задания на заполнение пропусков, флеш-карточки, вопрос-ответ.


Из плюсов можно отметить большое количество различных шаблонов и простоту их создания. Из минусов – большая часть шаблонов платная.

На рисунке 3 показана главная страница сайта на котором создаются печатные материалы.



Рисунок 3 – Окно программы QuickWorksheets

На рисунке 5 показан этап создания материалов. На данном этапе выбирается шрифт и размер шрифта, вводится заголовок страницы, так же пишется инструкция к тексту. Затем вводим текст. Во втором пункте нужно ввести ответы или подсказки. В остальных пунктах выбираем параметры страницы: строчные или заглавные буквы, количество колонок и т.д. В зависимости от шаблона. Пример созданного печатного материала в данном генераторе приведен на рисунке 4.



## Fruit and Vegetables

Name: \_\_\_\_\_  
 Date: \_\_\_\_\_

Unscramble the names of these common fruits and vegetables.

onion	watermelon	peach	pea	tomato
strawberry	radish	carrot	potato	zucchini
banana	pumpkin	lemon	grapefruit	broccoli
pear	grape	lime	apple	orange

1. elppa	_ _ _ i _	11. ridsah	_ a _ i _ _
2. tocrar	c _ _ _ o _	12. ioblocr	b _ _ c _ _ _
3. onnoi	_ n _ _ _	13. olime	_ _ m _ _
4. iugrtfeap	_ f a _ _ _ _ i _	14. rasyrerbwt	s t _ _ _ _ r _ _
5. ongear	_ _ _ n g _	15. ahepc	_ _ _ _ h
6. totaom	_ o _ _ t _	16. piunmkp	_ _ _ p _ _ n
7. ttpoa	p _ _ _ t _	17. nhociuzi	_ _ c c _ _ _
8. tnirecewam	_ a _ e _ m _ _ _	18. ape	_ e _
9. agpre	_ _ a _ _	19. ilme	i _ _ _
10. nbaaan	b _ n _ _ _	20. paer	_ _ a _

Created at [Quickworksheets.net](https://www.quickworksheets.net)
Page 1

Рисунок 4 – Созданные материалы

Aa Sans-serif

Aa Serif

Aa Hand

Size: Defau ▾



Enter Worksheet Title...

Name: \_\_\_\_\_

Date: \_\_\_\_\_

Enter instructions (optional)...

1. Words to Scramble:

Enter 1 word per line.

Use a Word List...

Words: 0 of 26

2. Answer Hints:

Pre-filled answer spaces:

☐ No hints

\_\_\_\_\_

☒ First letter filled

g \_ \_ \_ \_ \_

☐ 25% letters filled

\_ a \_ \_ \_ r \_ i

Word Bank:

☐ Show answers in a Word Bank.

radish

grape

lemon

apple

lime

grapefruit

pear

lemonade

strawberry

peach

watermelon

custard

orange

peach

lemon

lemon

pear

lemonade

juicyfruit

1. playe

...

9

11. slash

...

9

3. Real English Anagrams:

☐ Attempt to scramble into other words.

5. rapt figure

g r a p e f r u i t

Рисунок 5 – Создание печатных материалов

11

### 1.3 Программа WorksSheet Generator

Данный сайт предлагает 5 шаблонов для генерации материалов по английскому для проверки знаний учащихся. Главная страница сайта изображена на рисунке 6.

Создавать шаблоны на данном сайте достаточно просто. Окно в котором задаются параметры для листа с заданиями показано на рисунке 7. В поле 1 вводим заголовок рабочего листа. Во втором поле нужно ввести слова для левой колонки. В поле 3 вводим совпадающие элементы для правой колонки. В поле 4 нужно написать инструкцию, которая будет отображаться в верхней части листа. В поле 5 можно написать инструкцию, которая будет отображаться в нижней части листа. Данное поле заполняется по желанию. В поле 6 заполняется дополнительная информация, которая будет отображаться в нижней части листа с ответами.



Рисунок 6 – Главная страница WorksSheet Generator

**1** What is the **Title** of your worksheet?  
Opposite Adjectives

**2** What are the items for the left-hand column? Enter ten words or phrases.

hot
white
quiet
new
good
short
cheap
alive
right
clean

**3** What are the **matching** items for the right-hand column? Enter ten words or phrases.

cold
black
noisy
old
bad
long
expensive
dead
wrong
dirty

tip: We recommend that you fill all ten fields in each of 2 and 3. Leaving fields blank will produce unpredictable results.

**4** **Instructions at top of Worksheet:**  
Draw a line from each adjective on the left to its opposite adjective on the right.  
[What instructions do you want to give your students?]

**5** **Extra info at bottom of Worksheet (optional)**  
Insert a personalized reminder or tip for your students:  
[Double click inside field for drop-down list of suggestions, or enter your own text.]

**6** **Extra info at bottom of Answersheet (optional)**  
Insert a personalized note for your students:  
[Double click inside field for drop-down list of suggestions, or enter your own text.]

MAKE WORKSHEET →

MAKE ANSWERSHEET →

Clear Form

Reset Form

Рисунок 7 – Окно заполнения параметров для создаваемого листа с заданиями

**Matching Worksheet** – это шаблон, в котором нужно провести линии соответствия между парами слов или фраз. В двух столбцах имеются слова. Учащимся нужно провести линии между словами которые соответствуют друг другу. В шапке страницы также есть инструкция для правильного выполнения задания. Данный шаблон показан на рисунке 8.

## Opposite Adjectives

### Worksheet

**Instructions:** Draw a line from each adjective on the left to its opposite adjective on the right.

hot	wrong
white	noisy
quiet	old
new	dead
good	expensive
short	long
cheap	dirty
alive	bad
right	black
clean	cold

Made with the TEFL.net Worksheet Generator

EnglishClub.com

Рисунок 8 – Шаблон Matchin Worksheet

**Scrambled Words Worksheet** – в данном шаблоне студенты должны правильно расшифровать слова и записать их правильно. На странице в столбик записаны слова, а напротив слов имеются линии, куда нужно вписать правильный перевод слова. В шапке страницы также имеется инструкция к заданию. Данный шаблон показан на рисунке 9.

## Spelling - Animals

### Worksheet

**Instructions:** The animals below are scrambled. Write them out correctly.

1. atc \_\_\_\_\_
2. gdo \_\_\_\_\_
3. ocdceorli \_\_\_\_\_
4. onli \_\_\_\_\_
5. irteg \_\_\_\_\_
6. rnpheat \_\_\_\_\_
7. shroe \_\_\_\_\_
8. nitsec \_\_\_\_\_
9. dbir \_\_\_\_\_
10. isfh \_\_\_\_\_

Made with the [TEFL.net](http://TEFL.net) Worksheet Generator

[EnglishClub.com](http://EnglishClub.com)

Рисунок 9 – Шаблон Scrambled Words Worksheet

**Sorting Worksheet (2, 3 or 4 columns)** – в данном шаблоне студенты должны рассортировать слова в 2, 3 или 4 колонки. Внизу под колонками написаны слова, которые нужно вписать в соответствующие им колонки. В зависимости от шаблона количество колонок варьируется от 2-х до 4-х. В шапке страницы имеется инструкция к заданию. Шаблон представлен на рисунке 10.



## Countries & Continents

### Worksheet

**Instructions:** Sort the items from the box into the appropriate columns.

Africa	Asia	Europe	South America

Made with the TEFL.net Worksheet Generator

[EnglishClub.com](http://EnglishClub.com)

Brazil	Argentina	Germany	Angola	Romania	
Vietnam	UK	Uruguay	India	Austria	
Venezuela	Iran	Zambia	Chad	China	Japan
Libya	Madagascar	Finland	Chile		

Рисунок 10 – Шаблон Sorting Worksheet (4 columns)

### 1.4 Программа LearnClick

В программе LearnClick возможно создание различных вариантов упражнений для проверки знаний учащихся. Можно создать упражнения на заполнение пробелов в тексте (варианты можно выбирать из выпадающего меню или из таблицы над текстом). Так же есть возможность вставить в задание видео. Так же у преподавателя есть возможность выставить несколько



одновременно несколько правильных ответов. Варианты ответов выпадающего списка тоже можно вписать самостоятельно, иначе они будут сгенерированы из слов, которые решили скрыть.

Так же можно создавать задания на сопоставление слов и переводов.

Можно создавать тесты с одним или несколькими вариантами правильных ответов. Здесь же можно вставить комментарии для объяснения, которые будут показаны после прохождения теста.

Так же есть кнопка «я сдаюсь», ее можно вставлять в любое задание.

На данном сайте имеется возможность сохранять сгенерированные задания в файл формата pdf.

Есть возможность выставить ограничения на количество попыток прохождения тестов. Это может быть использовано при проведении контрольных работ.

Этап создания печатных материалов на данном ресурсе показан на рисунке 11.

Из минусов программы можно отметить платную подписку. Бесплатный доступ практически отсутствует, у пользователя есть 3 дня, а затем нужно решить, оформлять платную подписку или нет.

Ни одна из упомянутых программ не позволяет быстро и эффективно создать учебные тесты для работы с текстом в качестве предмета задания по иностранному языку.

Среди преимуществ разработанной программы «Генератор учебных заданий» можно выделить:

- простоту работы по генерации заданий,
- используется многомерный анализ данных для анализа текста на английском языке,
- генерация тестовых материалов, за счет привлечения средств автоматической генерации специальных шаблонов, ссылающихся на разделы многомерной базы данных.

Title:

Category:  [add a category](#)

Order	Results	Visible To
<input type="checkbox"/> One question per page <input type="checkbox"/> Randomize questions		
Max number of attempts: <input type="text" value="Unlimited"/>		Time Limit: <input type="text"/> Minutes

---

Question 1: Blank Boxes & Dropdowns ✖

Create Gap Remove Gap Add a Hint Make Dropdown More Options

**B** *I* U A Font size 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94 96 98 100

Can you pass me the milk, please ?

Explanation: (optional)

---

Question 2: Checkboxes (several answers correct) ✖

Create Gap Remove Gap Add a Hint More Options

**B** *I* U A Font size 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94 96 98 100

Which sentences are correct?

Options:

<input type="text" value="It will be rain soon."/>	<input type="checkbox"/> Correct Answer
<input type="text" value="It will be raining soon."/>	<input checked="" type="checkbox"/> Correct Answer
<input type="text" value="It is raining soon."/>	<input type="checkbox"/> Correct Answer
<input type="text" value="It is going to rain soon."/>	<input checked="" type="checkbox"/> Correct Answer

[Add an option](#)

Explanation: (optional)

---

Question 3: Drag & Drop ✖

Create Gap Remove Gap Add a Hint More Options

**B** *I* U A Font size 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94 96 98 100

Receptionist: — Good afternoon, Valentine Hotel. May I help you?  
 Mrs Smith: — Yes. I'd like to book a room, please.  
 Receptionist: — Sure. When for, madam?  
 Mrs Smith: — April the 15th.  
 Receptionist: — How long will you be staying?  
 Mrs Smith: — Three nights.  
 Receptionist: — What kind of room would you like?  
 Mrs Smith: — Double with bath.

Explanation: (optional)

---

Drag & Drop Add a question

User input only gets recorded if you choose 'LearnClick classes' under the settings tab 'Visible To'.

[Save](#)

[Show Quiz](#)

Рисунок 11 – Создание печатных материалов на LearnClick

## **2 Методы и средства разработки**

### **2.1 Порождающая грамматика над многомерными базами данных**

В данной работе рассматриваются принципы работы и задача разработки программы, позволяющей генерировать учебные задания, в целях упрощения и повышения качества работы по составлению учебных материалов.

Проблема генерации осмысленной речи исследуется достаточно давно. Проблема генерации осмысленной речи исследовалась различными авторами, такими как К. Шеннон, Т. Виноград, Э. Кодд, А. Хомский, М.В. Никитин, А.С. Нариньяни. Машинный перевод, перевод, а так же построение экспертных систем являются проблемами в данной области.

Разработка приложения состоит в повышении эффективности работы преподавателей, сокращении временных затрат, увеличении объемов генерируемых тестовых материалов, которые позволят повысить качество обучения за счет множества вариантов индивидуальных заданий. Для разработки обучающих ресурсов от преподавателя требуются привлечение значительных временных ресурсов, что затрудняет индивидуализацию траекторий обучения учащихся. Поэтому задача разработки ресурса, позволяющего преподавателям иностранных языков, генерировать учебные задания в целях проверки знаний учащихся, является актуальной. Таким образом, задача разработки данного ресурса является значимой, так как данный ресурс сможет помочь преподавателям уменьшить время создания электронных ресурсов и увеличить количество генерируемых материалов.

Приложение по генерации учебных заданий поможет преподавателям составлять индивидуальные задания для своевременной проверки знаний учащихся. Кроме того, разработанная система позволит экономить время работы преподавателя. Большое количество генерируемых материалов позволит предотвратить списывание ответов недобросовестными учениками на уроке и повысит объективность оценки знаний студентов.

Разработка программы направлена на гибридизацию моделей баз данных и порождаемых шаблонов. Порождаемые шаблоны помогут сделать задания уникальными, а так же разнообразить их содержание. Использование генераторов учебных заданий будет более привлекательно, чем любое незаконное использование или редактирование чужих материалов с нарушением авторских прав.

Порождаемые шаблоны – шаблоны построения фраз естественного языка, которые могут быть заданы, как вручную, так и порождаться на основе, в частности, контекстно-свободных порождающих грамматик [45].

Они позволяют сделать систему независимой от способа создания, композиции и представления объектов.

Action + -ing + Link + -(e)s + Process + (Manner);

Action + -ing + the + Object + Link + -(e)s + Process + (Manner);

Action + -ing + was + Done + Group of Time;

Surely, + Action + -ing + was + Done + Group of Time [45].

Маркеры типа Action, Object обращаются к соответствующим ячейкам или группам ячеек условно трехмерного пространства данных с учетом текущей темы генерации учебных заданий [45-50].

Порождающие грамматики, помогут осуществить порождение шаблонов генерации, которые будут ссылаться на ячейки многомерных баз данных. На основе шаблонов, которые поможет сгенерировать порождающая грамматика, программа генератор заданий генерирует фразы, имеющие вид:

This responsible producer invents various keyboards. This responsible user sets up a complex computer. That clever woman makes a right project. This clever researcher purchases a large desktop.

Порождающая грамматика Хомского задается как множество правил порождения, где каждая пара цепочек является правилом и дает возможность заменить левую цепочку на правую.

Из терминальных и нетерминальных символов могут быть составлены цепочки в правилах грамматики. Терминальный алфавит обозначают

символом  $T$ . В промежуточных порождениях цепочек используют нетерминальные символы. Нетерминальный символ обозначается как  $S$  и обозначает все цепочки языка. В каждой порождающей грамматике левая часть должна состоять из единственного начального нетерминала, иначе нельзя будет породить ни одной цепочки.

При помощи определенных правил подстановки, порождающая грамматика позволяет порождать цепочки языка из некоторых начальных цепочек.

Порождающая грамматика имеет вид упорядоченной четверки:  $G = (S, T, N, R)$ , здесь  $T$  – терминальный алфавит,  $N$  – множество нетерминальных символов,  $S$  – начальный символ порождающей грамматики,  $R$  – конечное множество правил порождения [45-47].

Порождение – это процесс в котором на каждом шаге из цепочки, уже полученной на предыдущем шаге (в частности из начальной), с помощью применения к ней правил замены, можно получить новую цепочку [45-47].

### **Типы грамматик.**

Хомским были введены классы грамматик задав ограничения на вид правил порождающей грамматики [9-10].

#### **Тип 0 (неограниченные).**

Грамматикой типа 0 является любая порождающая грамматика, без каких либо ограничений. Никакие дополнительные ограничения на данный тип правил грамматики не накладываются [9-10].

#### **Тип 1 (контекстно-зависимые).**

Частный случай формальной грамматики, у которой левые и правые части могут быть окружены терминальными и нетерминальными символами.

#### **Тип 2 (контекстно-свободные).**

Частный случай формальной грамматики, у которой левые части всех продукций являются одиночными нетерминалами (объектами, обозначающими какую-либо сущность языка (например: формула, арифметическое выражение, команда) и не имеющими конкретного символического значения). Смысл термина

«контекстно-свободная» заключается в том, что есть возможность применить продукцию к нетерминалу, в отличие от общего случая неограниченной грамматики Хомского, не зависит от контекста этого нетерминала [9-10].

### **Тип 3 (Регулярные).**

Регулярные грамматики определяют в точности все регулярные языки, и поэтому эквивалентны конечным автоматам и регулярным выражениям. Регулярные грамматики являются подмножеством контекстно-свободных [9-10].

### **Дополнительный тип.**

В диссертационной работе был разработан дополнительный тип порождающих грамматик на основе гибридизации контекстно-свободных порождающих грамматик и ссылок на массивы многомерных баз данных: порождающие грамматики над многомерными базами данных. Данный тип помогает решить проблему создания многовариантных учебных заданий, для качественной проверки знаний учащихся.

В синтаксис порождающих грамматик вписываются координаты многомерных данных. Огромные массивы информации выносятся из наполнения правил порождающих грамматик в массивы слов многомерной базы данных.

При порождении фраз порождающей грамматикой используются ссылки на файлы с группами слов, названия которых содержат в себе указание координат многомерного пространства данных, что полностью воспроизводит и заменяет собой СУБД для решения таких простых задач как генерация учебных заданий в рамках конкретной тематики рассматриваемых текстов.

## **2.2 OLAP технологии и применение многомерных баз данных при генерации осмысленной речи**

OLAP – технология обработки данных, заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу.

В реляционных базах данные хранятся в отдельных, хорошо нормализованных таблицах. Для операционных баз данных эта структура удобна, но сложные многотабличные запросы обрабатываются в ней медленно. Поэтому, чтобы получить скорость для обработки запросов, используют OLAP технологию.

Существует три типа OLAP:

1. MOLAP – многомерная OLAP. Данный тип является классической формой OLAP. Здесь данные хранятся в многомерных структурах и за счет этого повышается скорость обработки запросов [26].

2. ROLAP – реляционная OLAP. Здесь хранение данных происходит в реляционных таблицах. Для хранения агрегатов создаются дополнительные реляционные таблицы. Реляционные таблицы помогают хранить большое количество данных. В реляционной OLAP многомерные данные можно представить с помощью двух схем, это схема «звезда» и схема «снежинка» [26].

Плюс данного типа: все данные хранятся в одной базе и в одном формате. Минус: значительно увеличивается объем таблиц данных.

3. HOLAP – гибридная OLAP. Данный тип сочетает MOLAP и ROLAP, за счет чего минимизирует недостатки и использует достоинства обеих структур. Здесь для хранения данных используются реляционные таблицы, а для хранения агрегатов используются многомерные таблицы. Данный тип значительно уступает по скорости MOLAP [26].

Для удобства просмотра и анализа информации используется OLAP куб. В кубе данные хранятся в одном месте, поэтому к данным могут обращаться одновременно много пользователей, независимо от того где находятся.

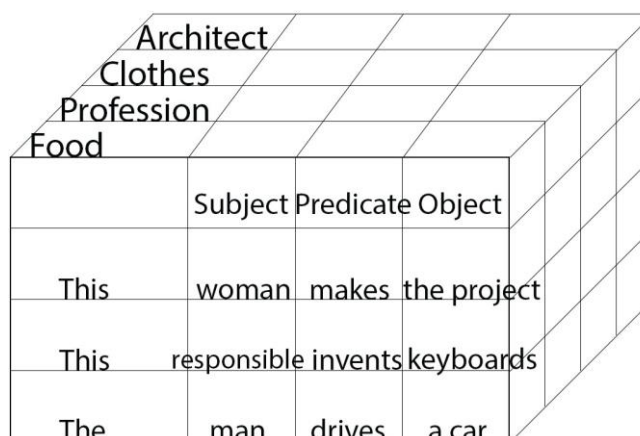


Рисунок 12 – Пример трехмерного куба

На рисунке 12 показан пример трехмерного куба. В качестве измерений в трехмерном кубе используются темы, а в качестве мер используются сгенерированные предложения.

Для визуализации информации на компьютере трехмерный куб не удобен, поэтому используют двумерные представления данных, или по-другому, табличные. Для того чтобы получить двумерный куб, его необходимо разрезать поперек одной или несколькими осями.

Architect	Builder	Brick	Building
Clothes	Seamstress	Checked	Dress
Profession	Programmer	Powerful	Computer
Food	Cook	Appetizing	Pizza

Рисунок 13 – Двумерный срез трехмерного куба



На рисунке 13 представлен двумерный срез куба. В заголовках столбцов представлено одно измерение, в заголовках строк другое измерение, а в ячейках таблицы представлены значения мер.

На рисунке 14 приведен пример базы данных, которая используется в программе и сделанная с помощью Microsoft Access.

Код	Поле1	Поле2
21	of_course	конечно
22	in_fact	фактически
23	maybe	может_быть
24	perhaps	возможно
25	in_principle	в_принципе
26	practically	практически
27	generally	в_общем
28	in_particular	в_частности
29	for_example	например
30	for_instance	например
31	moreover	тем_более_что
32	besides	кроме_того
33	furthermore	кроме_того
34	at_the_same_time	в_тоже_самое_время
35	therefore	из-за_этого
36	as_for_me	что_касается_меня
37	well	ну
38	thus	таким_образом
39	so	так
40	in_this_way	таким_образом

Рисунок 14 – Пример базы данных

### 2.3 Использование метрики Хэмминга для оценки близости генерируемых заданий к тексту

Расстояние Хэмминга – это расстояние между двумя точками многомерного, обычно символьного, пространства данных, заданной параметром длины, равным количеству различных координат вектора положения каждой точки.

$$DemCop < \sum_{i=1}^k HD(j|HD \rightarrow \min) ((w[a'_1, a'_2 \dots a'_k]), (v[a''_1, a''_2 \dots a''_k])) < RelText \quad (1)$$

где DemCop – требования к копирайту;

RelText – релевантность фраз к тексту;

HD – расстояние Хэмминга;

W – слово, построенное на многомерной базе данных, которое отображает фразу из текста;

V – вектор признака для каждой структуры, которая порождается.

Как известно пространство определяется через несколько базовых понятий, в частности множество точек, множество подточек с заданными на нем расстояниями между точками, Один из бесконечного числа возможных базисов пространства, то есть системы координат исходящей из точки начала координат. Кроме того на любом пространстве определяется метрика, то есть функция вывода расстояния AC на основе двух аргументов расстояния AB и расстояния BC.

Традиционная метрика является Евклидовой, то есть данная функция является корнем квадратным от суммы квадратов. Известны так же другие метрики. Одним из наиболее простых случаев метрики, является метрика Хэмминга, которая определяется следующим образом: пусть A – это точка с координатами  $[a_1, a_2, \dots, a_n]$ , B  $[b_1, b_2, \dots, b_n]$ , тогда расстояние AB будет равняться количеству координат  $a[i], b[i]$ , где  $a[i]$  не равняется  $b[i]$ .

Таким образом, метрика Хэмминга наиболее подходит для описания координат в многомерном пространстве для семантических объектов часто выраженных строками символов. Применительно к лексико-семантическому пространству просматриваемому в данной работе необходимо правильно выделить систему координат. В этом случае мы рассматриваем не четырехмерное пространство с координатами: порядок слов, варианты, темы, частотность, а рассматриваем более частные координаты, такие как группа тем, отличительный признак подгруппы тем, отличительный признак более частной

темы и т.д. В этом случае метрика Хэмминга дает хорошую качественную оценку семантической близости слов и групп слов.

Приведем пример использования метрики Хэмминга для реализации практических целей нашей работы. Компьютер мощный и медленный, и компьютер портативный и настольный. Расстояние Хэмминга между словами «портативный» и «настольный» будет меньше, чем расстояние между словами «портативный» и «мощный».

Данный метод дает возможность генерировать фразы с достаточным уровнем отклонения от исходного текста.

## **2.4 Семантический шум**

Семантический шум представляет собой модель одного из проявлений естественного языка – творческого иносказания и трансформаций фраз человеком на основе определенного диапазона возможностей языка. Формально отвлекаясь от типов связей внутри фраз с семантическим шумом можно породить одну базовую интерпретацию текста на основе порядка терминов в многомерной семантической базе данных. Множество фраз не приведенного вида может быть получено из данной формы высказываний (функций над ячейками многомерной базы данных), полученные высказывания представляют собой иносказания стандартных фраз более или менее высокого художественного достоинства. На основе таких фраз строится множество учебных заданий, интегрируемых в различные варианты тестов.

Очевидно, что само смысловое значение, включающее все ассоциативные и эмоциональные компоненты, не может быть формализовано на основе простого вектора рационально представленных иррационально-смысловых категорий. Но в рамках определенных аппроксимаций и приближений можно описывать семантику естественного языка до уровня достаточного для решения конкретных инженерных задач разной степени сложности. В перспективе на основе семантического понятийного аппарата имеется принципиальная, хотя и

чрезвычайно сложная, задача прохождения теста Тьюринга лингвистическим программным обеспечением. При этом остается актуальным разработка программных систем с более узкой функциональной направленностью, в частности, рассматриваемый программный продукт.

В рамках такого рода работы над узким набором задач необходимо развивать принципы генерации осмысленной речи на основе проработки отдельных подмножеств многомерной классификации слов и понятий языка с учетом возможности использования этих фрагментов для генерации подмножеств естественного языка. Необходима реализация этих принципов на основе генерируемых (порождающими грамматиками) семантических шаблонов с извлечением из их структуры семантического шума с частичными семантическими потерями.

Порождение шаблонов генерации осмысленной речи, ссылающихся на ячейки многомерной базы данных, может быть осуществлено при помощи порождающих грамматик и впоследствии уточняться на основе введения операций над векторами сем, когда будут вводиться запреты на подмножества комбинаций фраз, генерируемых на основе принципа семантических подстановок. Данная программа может позволить составлять различные варианты заданий одного типа. Множества вариантов генерируемых шаблонов позволят индивидуализировать задания, так как большое количество тестовых материалов необходимо для исключения списывания со стороны студентов и учеников. Также она позволит обеспечить экономию времени на разработку образовательных ресурсов преподавателям.

В качестве одного из этапов в направлении интерпретации текста на естественном языке необходимо осуществить проекцию слов каждого предложения такого текста с учетом реализуемой (в качестве побочного) эффекта экстракции семантического шума. На этом этапе необходимо приведение фраз текста к фразам приведенного вида с определенной вероятностью искажения смысла этих фраз, что технически не влияет на процесс генерации учебных заданий.

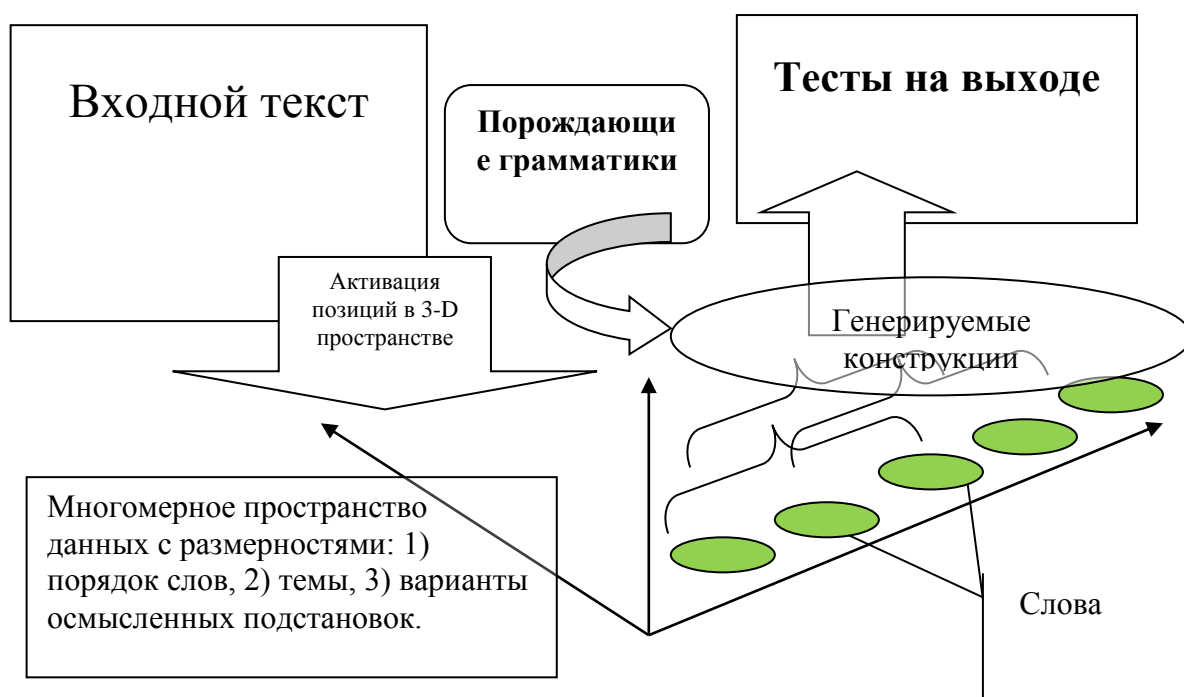


Рисунок 15 – Схема генерации осмысленных учебных заданий к текстам

## 2.5 Математическая модель

Генерируемое дерево состояний строки на основе порождающих грамматик с правилами, ссылающимися на разделы многомерных баз данных, реализуется в программной системе генератор порождаемых шаблонов позволяющей получать на выходе для конечного пользователя учебные материалы следующего вида.

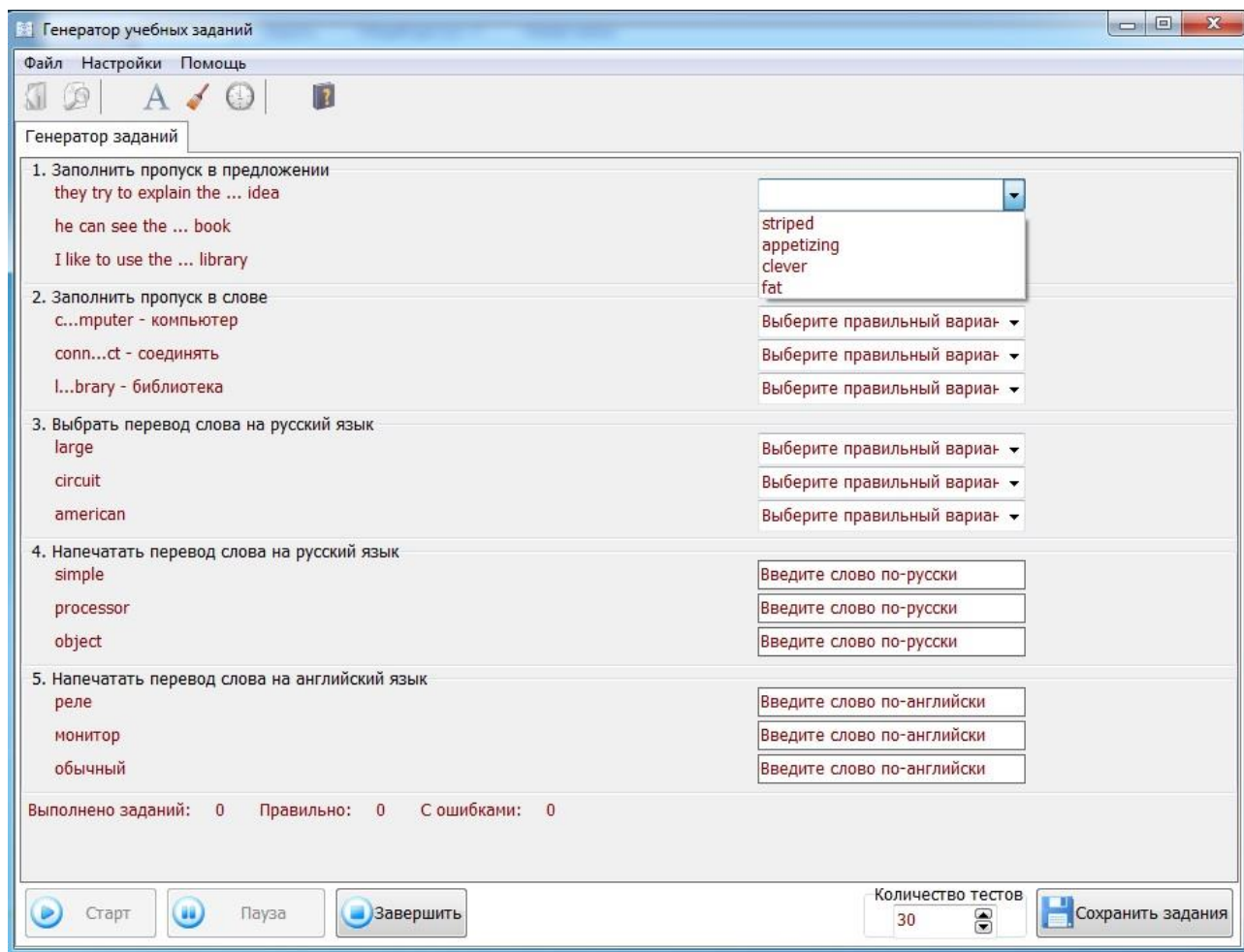


Рисунок 16 – Учебные тесты в программе «Генератор учебных заданий»

Эта модель данных дает возможность реализовать функционал разрабатываемой программной системы на основе алгоритмов, в основе которых лежат представления о системе естественного языка.

Большинство авторов анализируют язык с точки зрения грамматики, что делает затруднительным возможность составлять осмысленные предложения в необходимом для процесса обучения количестве.

Эта модель данных является отражением семантической структуры представления о естественном языке, что обеспечивает возможность выделения множества осмысленных фраз из множества бессмысленных, на основаниях критериев осмысленности, в частности в работах таких ученых семасиологов, как Никитин описываются понятия импликации одной группы слов на основе

другой. Так, например, слово здание может имплицировать следующее слово кирпичное, бетонное, высокоэтажное, жилое, комфортабельное. При наличии не двух, а нескольких таких групп слов накладываются ограничения на множество осмысленных, частотных, естественно звучащих фраз порождаемых на основе таких соответствий.

Программная система учитывает частотные принципы организации подмножеств естественного языка. В частности, разработанная программная система содержит в себе различные клоны одних и тех же групп слов в зависимости от их частотности. В результате чего порождаемые фразы имеют вид «Этот странный повар варит осетрину» в отличие от фраз, генерируемых без учета частотных принципов «Мой друг съел соленые огурчики» (частотная), «Разжигатель войны чавкает чрезвычайно размороженной пиццей» (не частотная).

В результате мы решаем проблемы выделения множества осмысленных фраз, из множества любых фраз включая бессмысленные, но и выделение из множества осмысленных фраз наиболее частотных.

Многомерный анализ текста в программе генератор учебных заданий дает возможность учитывать прецеденты употребления сочетаний слов в текстах, по мотивам которых создаются учебные задания.

Генерируемое дерево состояний строки на основе порождающих грамматик с правилами, ссылающимися на разделы многомерных баз данных, является элементом новизны данного диссертационного исследования. С ее помощью реализованы алгоритмы использующие критерии осмысленности и благозвучности порождаемых фраз разного вида.

Множество данных соответствующее этой математической модели не набирается в ручную, а копируется либо полуавтоматически на основе программных продуктов, являющихся бесплатными и реализуемыми в качестве инициативно открытого проекта СФУ, либо с их же помощью, но с использованием «mouse click» агента. В результате пользователь может получить очень большое множество генерируемых материалов

обеспечивающих потребности разных групп учащихся, отличающихся разным уровнем владения английским языком, разными интересами, набором мотивирующих тем, их психологическим особенностям. При помощи «mouse click» агента и привлечения других программ можно озвучить данные материалы. Имеется возможность при помощи «mouse click» агента возможность поиска картинок к генерируемому тексту, что особенно важно для студентов с аудиальным и визуальным складом соответственно. Для студентов с кинестетическим складом восприятия к данным учебным материалам можно сгенерировать большое количество учебных заданий на основе программы «patterns maker» используя все тот же «mouse click» агент.

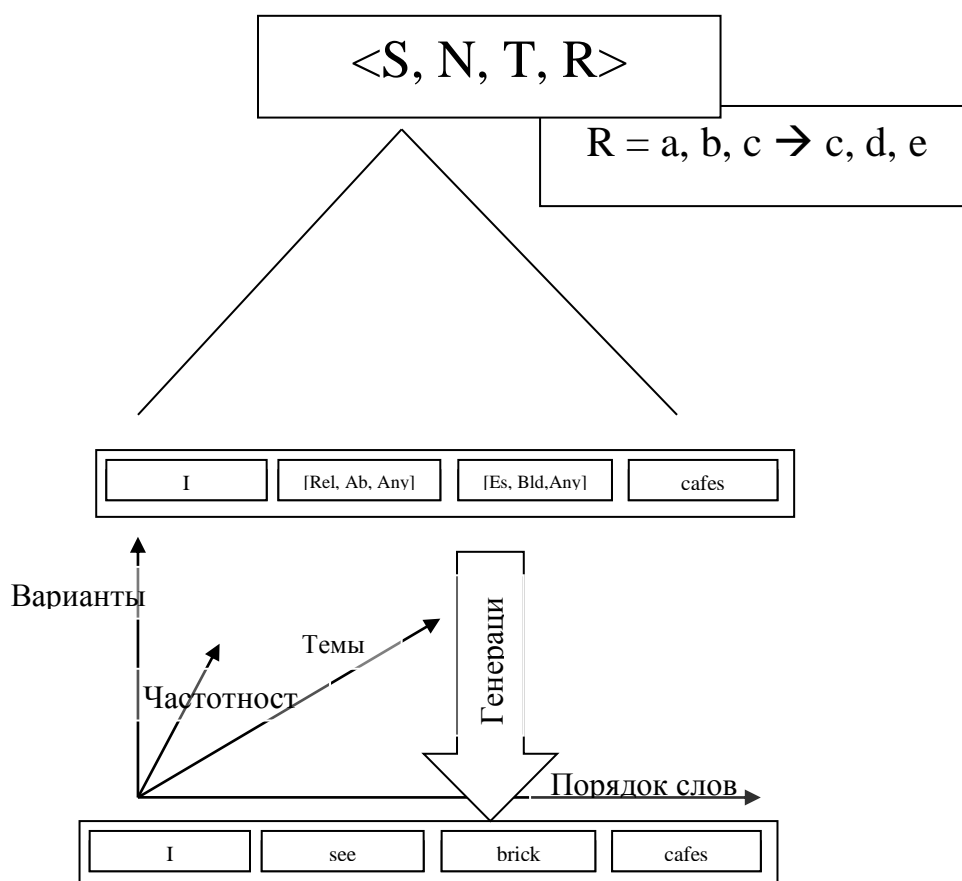


Рисунок 17. Схема генерации предложений рассматриваемой модели



### 3 Описание используемых программ в среде mouse auto clicker

#### 3.1 Описание программы «Electronic Dictionary»

Работа большинства программ строиться на основе словарей. Одной из ключевых программ для генерации таких словарей и подстановочных таблиц служит «Electronic Dictionary». На рисунке 18 представлен интерфейс программы электронный словарь «Electronic Dictionary 2010». Данная программа представляет собой набор программ, которые помогают изучать английский язык. В программе можно ознакомиться с универсальной классификацией слов английского языка, в основе этой классификации лежит иерархия понятий, более общие понятия находятся на вершине, более частные на низких узлах.

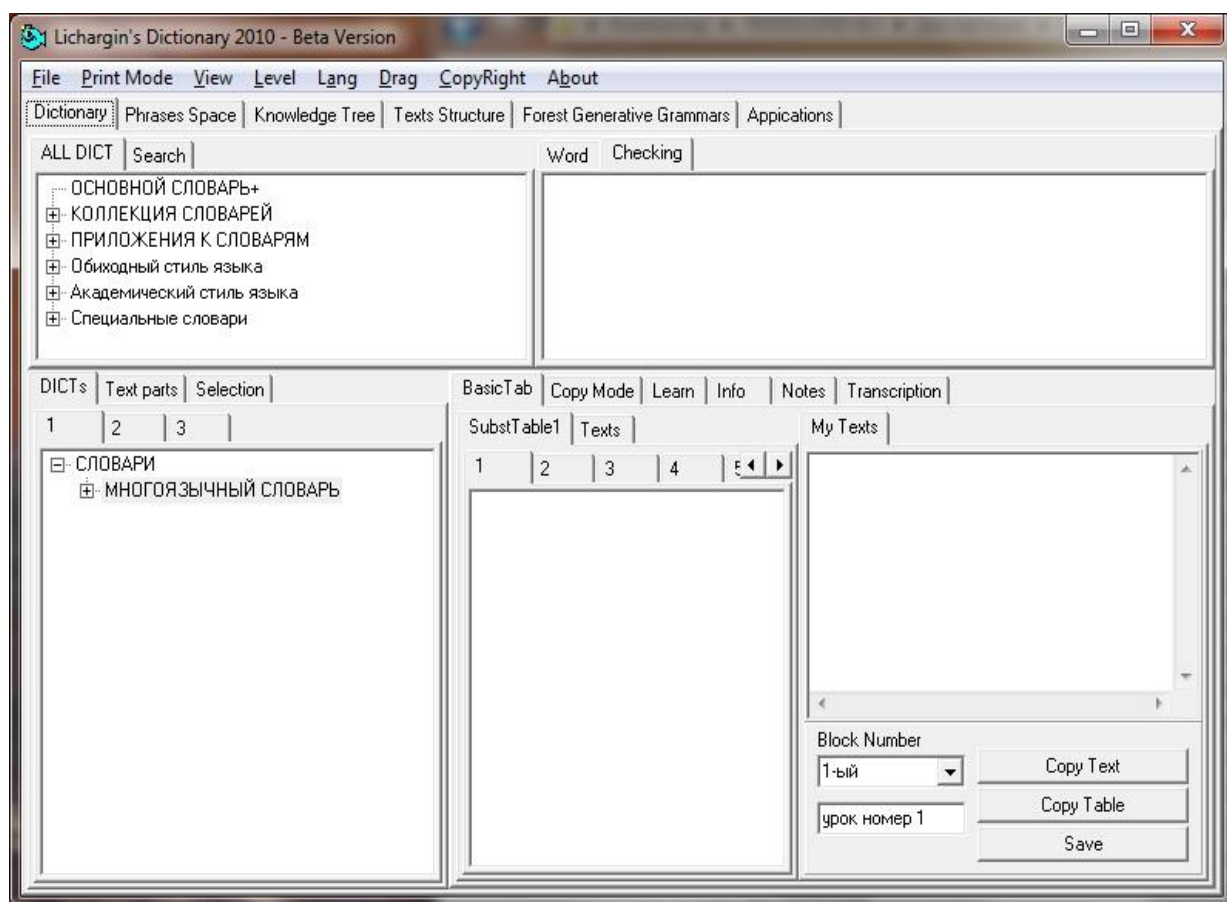


Рисунок 18 – Основное меню программы

В окне «ALL DICT» представлен основной словарь.

В закладке «Search» можно ввести слово, которое необходимо найти в словаре. Данное окно представлено на рисунке 19. В строку вводим слово, затем нажимаем кнопку «Start» и в окне выведется искомое слово. Можно посмотреть как его перевод, так и транскрипцию данного слова.

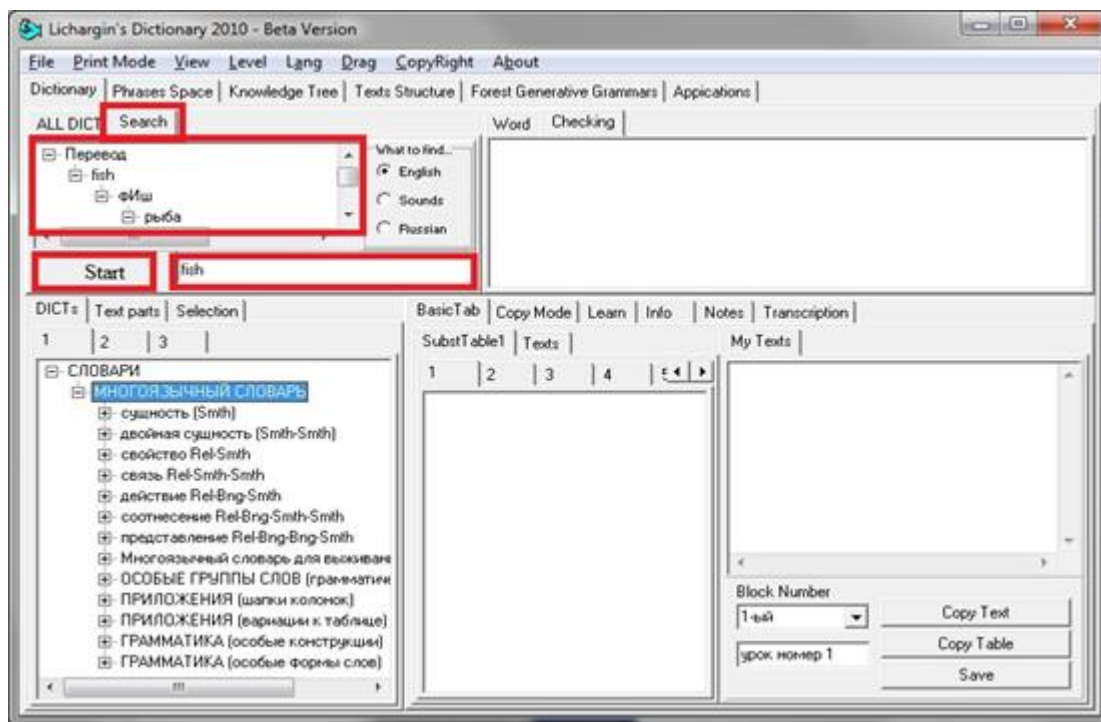


Рисунок 19 – Поиск слова в словаре

Так же в данной программе можно слова добавлять на изучение и потренировать слова. Для этого нужно зайти на вкладку «Learn». Данное окно представлено на рисунке 20. Для тренировки нужно выбрать слово, затем в выпадающем меню выбрать предполагаемый правильный вариант, а затем нажать кнопку «Check me». Ниже в графе «Results» выведется ответ, верно или неверно решено задание.

Перед тренировкой слов, необходимо добавить их на изучение. Для того чтобы добавить слова, нужно в словаре выбрать раздел, который будет изучаться, в соседнем окне в закладке «Basic tab» появятся слова которые имеются в данном разделе. Нажав правой кнопкой в окне нужно выбрать пункт

из выпадающего меню «Take for training more» и слова появятся на изучении в разделе «Learn». Данный этап изображен на рисунке 21.

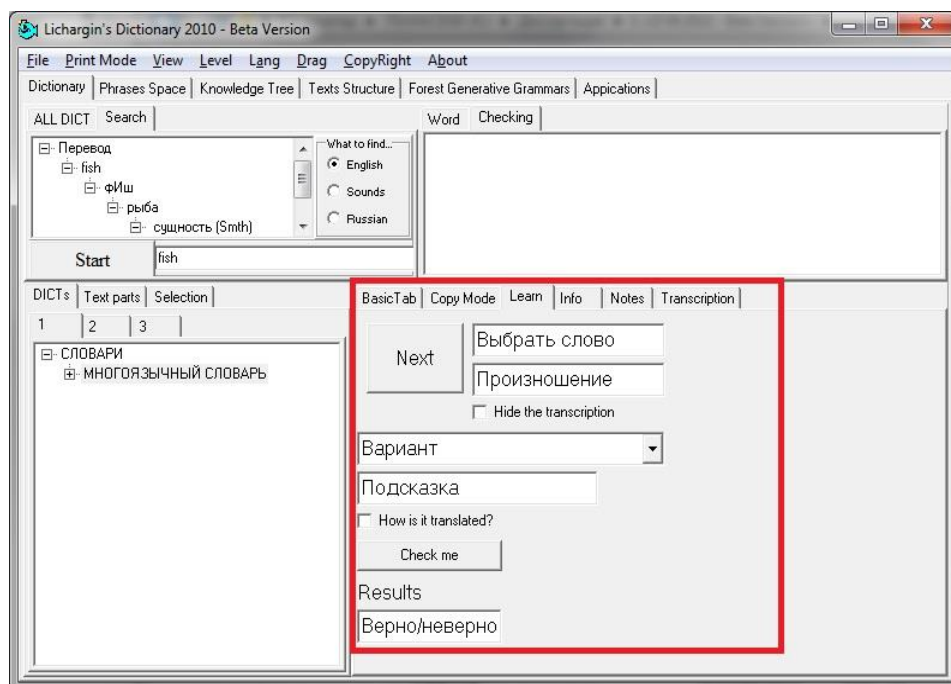


Рисунок 20 – Тренировка слов

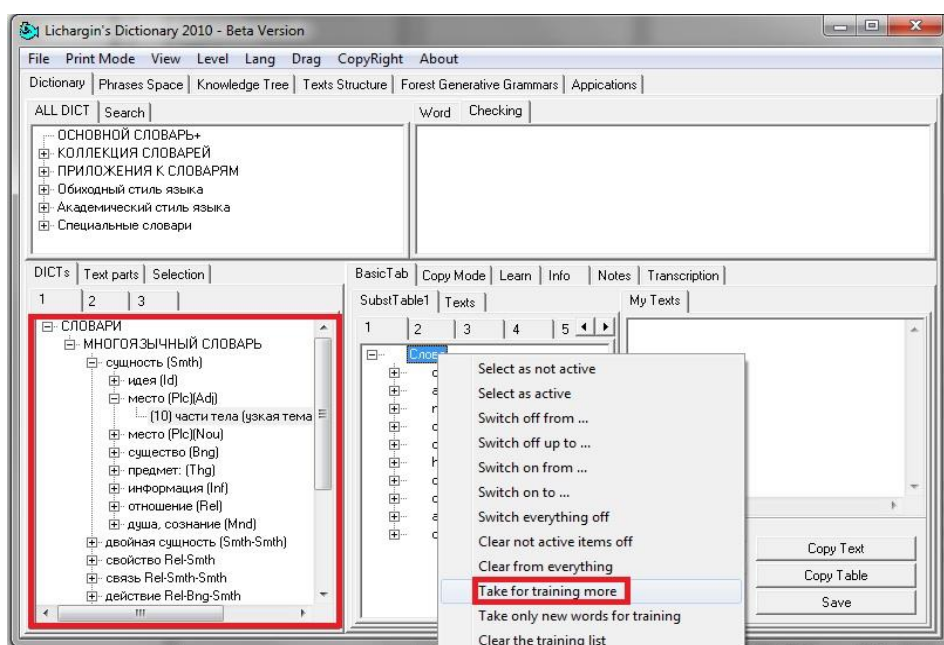


Рисунок 21 – Добавление слов для тренировки

Ниже на рисунке 22 показана возможность выбора уровня сложности. На выбор предлагаются 4 уровня сложности: Minimal, Beginner, Professional, Minus Beginner.

Minimal означает, что у пользователя минимальные знания, то есть владеет небольшим запасом элементарных слов.

Beginner — это начальный уровень владения английским языком. То есть предполагается, что у пользователя есть небольшой словарный запас, а так же он может составлять элементарные предложения и фразы.

Professional предполагает, что пользователь отлично владеет английским языком. Свободно говорит и понимает на английском, а так же имеет большой словарный запас.

Minus Beginner означает, что элементарные слова в данном уровне не используются и предполагает, что у пользователя имеется хороший словарный запас и знания английского языка.

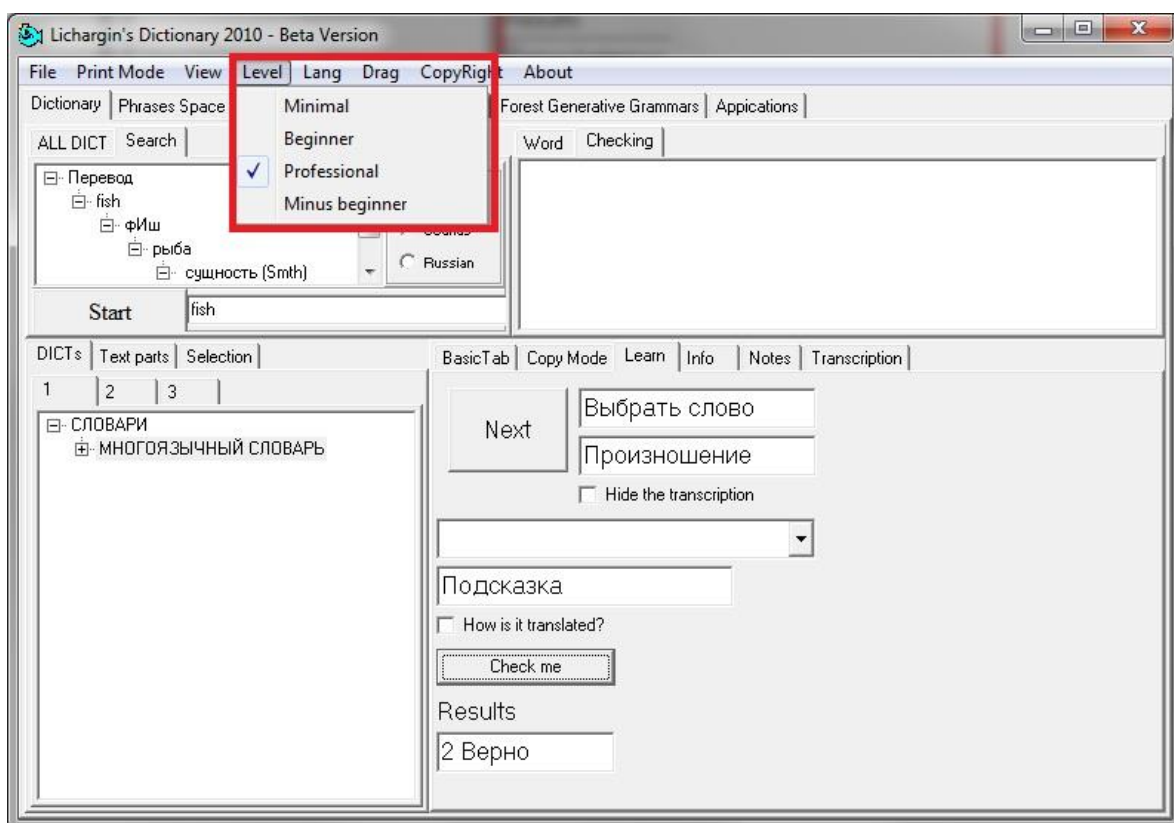


Рисунок 22 – Выбор уровня сложности

Так же в программе можно сохранить слова в файл. На панели меню в закладке «Print Mode» из выпадающего списка можно выбрать параметры сохранения. Можно сохранить слова в три колонки, в линию, включить или отключить транскрипцию, включить перевод с русского на английский или с английского на русский. Так же выбрать тип файла .html или .txt. На рисунке 23 изображено данное окно с выбором параметров сохранения в файл.

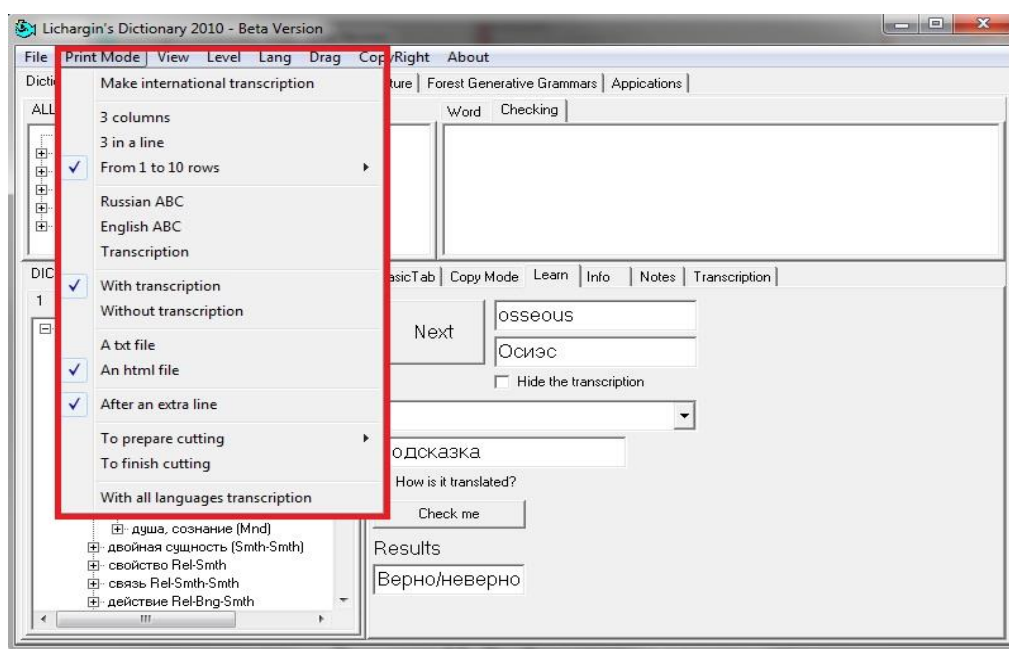


Рисунок 23 – Выбор шаблона для сохранения в файл

corporal	кОпэрэл	телесный
aural	О:рэл	ушной
nasal	нЕйзэл	носовой
cerebral	сЕрибрэл	мозговой
dental	дЕнтэл	зубной
head	хЕд	головной
cutaneous	кйу:тЕниэс	кожный
osseous	Осиэс	костяной
anal	Ейнэл	анальный
oral	О:рэл	ротовой

Рисунок 24 – Сохраненный файл

На рисунке 24 изображен пример сохраненного файла. В файле имеется три колонки, в первой слово на английском, во второй колонке транскрипция и в третьей колонке перевод слова, так как для сохранения были выбраны параметры: «3 columns», что означает сохранение в три колонки. Так же был выбран параметр «With transcription», что означает сохранение с транскрипцией. Формат сохраненного файла был выбран .html.

### **3.2 Описание программы «Acapela»**

Для создания аудио материалов используется программа «acapela», находящаяся по адресу [www.acapela-group.com](http://www.acapela-group.com). Главная страница программы представлена на рисунке 25. На рисунке видно, что для начала работы необходимо выбрать:

1. выбрать язык озвучиваемого текста;
2. выбрать голос;
3. ввести в 3 поле текст;
4. нажать кнопку «listen!».



Рисунок 25 – Озвучивание текста



Рисунок 26 – Расширения для Firefox

В браузере Firefox устанавливаем свободное расширение для Firefox, для



обработки одной и нескольких ("все" и "Выбор") загрузок с нескольких внешних менеджеров FlashGot!, находящееся по адресу <https://flashgot.net/>. Для его применения нужно перейти по вкладке Инструменты - FlashGot – Мультимедиа. Данный этап изображен на рисунке 26.

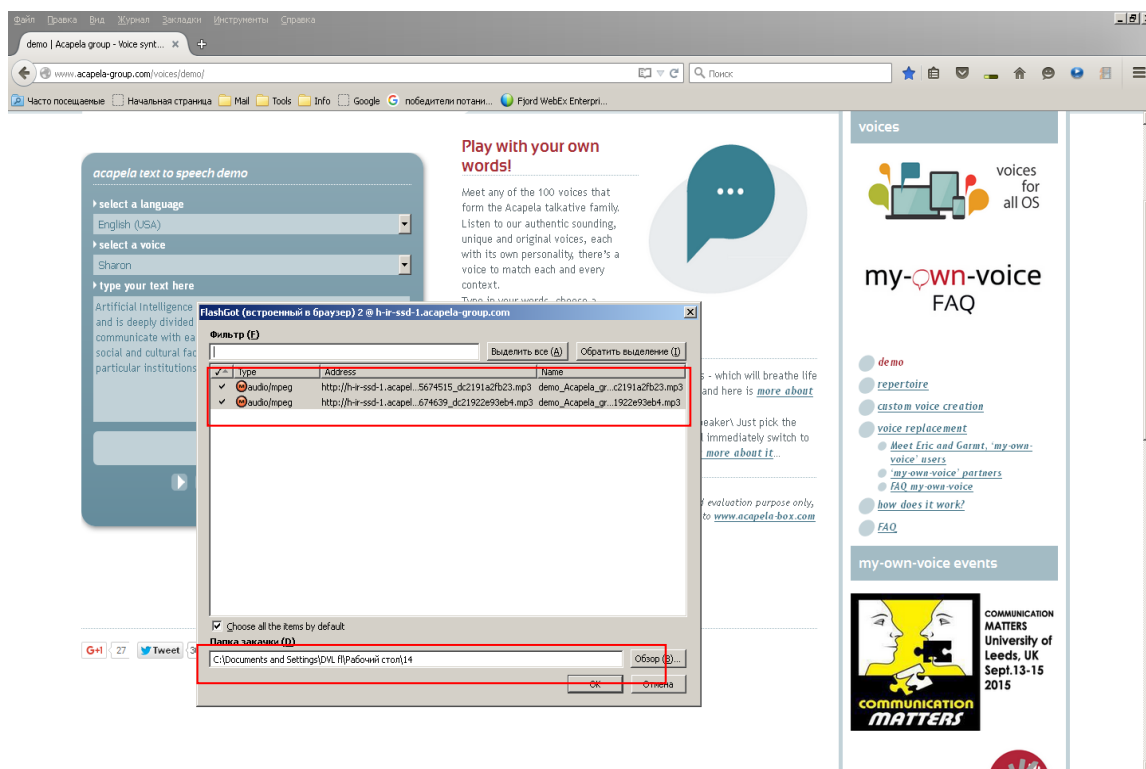


Рисунок 27 – Сохранение аудио файлов

Чтобы сохранить звуковые файлы, нужно в появившемся окне нужно отметить звуковые дорожки, выбрать папку для закидывания аудио файлов и нажать кнопку «ОК», после чего будет скачан файл. Данный этап изображен на рисунке 27.

### 3.3 О системе электронного обучения Сибирского федерального университета

В соответствии с новым законом об образовании в Сибирском федеральном университете широко практикуется использование электронного



обучения, дистанционных образовательных технологий, что уже отражается в документации к основным образовательным программам и в учебных программах дисциплин.

Разработка ЭОР высокого качества подразумевает большую учебно-методическую работу. Содержание ЭОР и применяемые инструменты обучающих систем могут замещать часть аудиторных занятий. Преподаватель, при организации аудиторных занятий, может планировать работу с учетом того, что студенты часть нагрузки выполняют с помощью ЭОР. В свою очередь, ЭОР могут избавить преподавателя и студента от рутинных действий (выдача заданий, проверка работ, комментирование и рецензирование, доступ к справочным и дополнительным материалам).

Кроме того, с помощью ЭОР можно повысить уровень организации самостоятельной работы студентов.

В 2013/14 учебном году ЭОР размещались, как правило, преподавателями-разработчиками учебных материалов на сетевых ресурсах [ms.sfu-kras.ru](http://ms.sfu-kras.ru), [study.sfu-kras.ru](http://study.sfu-kras.ru). Однако согласно приказу ректора № 748 от 17 июня 2014 года с нового учебного года электронные курсы СФУ будут размещаться в новой системе, по адресу [e.sfu-kras.ru](http://e.sfu-kras.ru). Таким образом, возникает задача переноса образовательного контента, в том числе мультимедийного, в новую систему.

Новая система представляет собой ресурс, разработанный на платформе Moodle, и имеет следующий вид, представленный на рисунке 28.

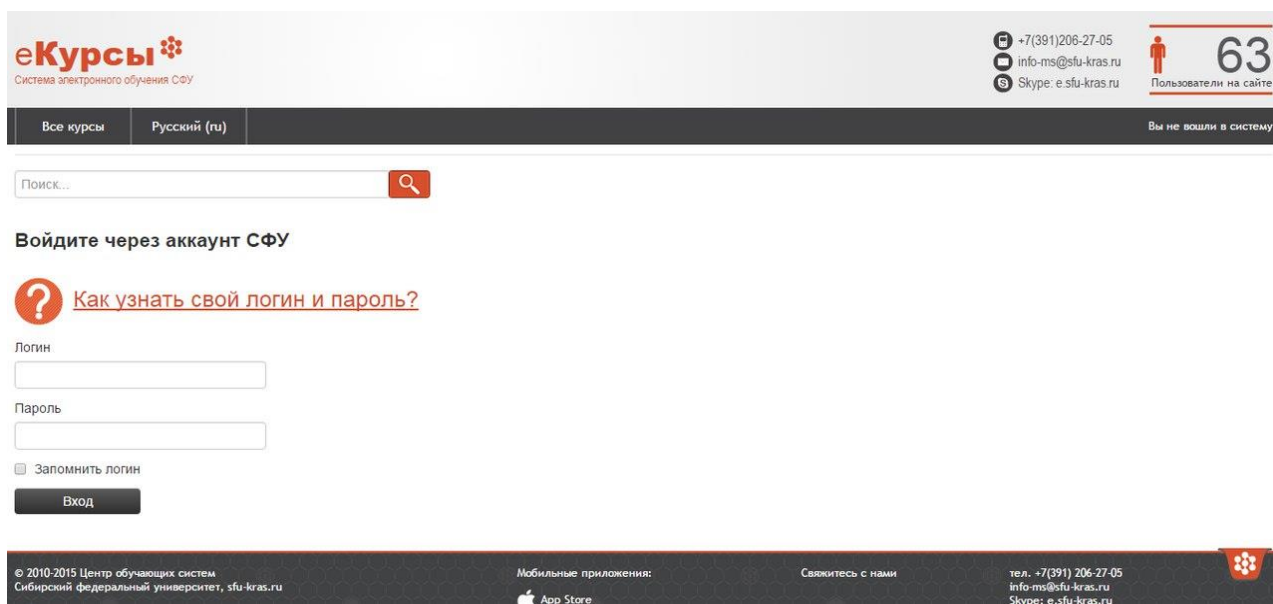


Рисунок 28 – Система электронного обучения СФУ на платформе Moodle

Moodle – система управления содержимым сайта, специально разработанная для создания онлайн-курсов преподавателями. Представляет собой инструментальную среду для разработки как отдельных онлайн-курсов, так и образовательных веб-сайтов. В основу проекта положена теория социального конструктивизма и её использование для обучения.

Moodle является бесплатно распространяемым программным комплексом (распространяющееся по лицензии GNU GPL), который по своим функциональным возможностям, простоте освоения и удобству использования удовлетворяет большинству требований, предъявляемых пользователями к системам электронного обучения.

Система ориентирована прежде всего на организацию взаимодействия между преподавателем и учениками, хотя подходит и для организации традиционных дистанционных курсов, а также поддержки очного обучения.

Moodle переведена на десятки языков, в том числе русский, и используется почти в 50 тысячах организаций из более чем 200 стран мира. В РФ зарегистрировано более 1000 инсталляций.

В России официальным партнером Moodle является ООО «Открытые технологии».

Moodle написана на PHP с использованием SQL-базы данных (MySQL, PostgreSQL, Microsoft SQL Server и др. БД — используется ADOdb XML). Moodle может работать с объектами SCO и отвечает стандарту SCORM.

Благодаря развитой модульной архитектуре, возможности Moodle могут легко расширяться сторонними разработчиками. Существенное расширение функциональных возможностей Moodle достигается за счёт интеграции подсистемы для организации вебинаров/веб-конференций. Помимо языковой поддержки и шаблонов оформления, Moodle позволяет подключать также следующие типы модулей: Элементы курса, отчеты администратора, типы заданий, плагины аутентификации, блоки, форматы курсов, отчеты по курсам, поля базы данных (для элемента курса «База данных»), плагины подписки на курсы, фильтры, отчеты по оценкам, форматы экспорта оценок, форматы импорта оценок, портфолио, типы вопросов в тестах, форматы импорта/экспорта тестов, отчеты по тестам, хранилища файлов, типы ресурсов, плагины поиска.

### **3.4 Описание работы программы «Mouse auto clicker»**

Mouse auto clicker – это программный продукт, который помогает оптимизировать работу пользователя. Программа имитирует различные действия пользователя. С помощью данного инструмента пользователь сможет записать действия и задать промежуток времени, через который будут выполняться эти действия, тем самым оптимизировав свою работу и сэкономив время.

При работе с программой пользователю необходимо записать свои действия, настроить скорость записанных действий, а затем запустить записанную последовательность. Так же можно выбрать действие, которое будет выполнено после воспроизведения записанной последовательности:

бездействовать, выход из системы, блокировать компьютер, включить режим сна.



Рисунок 29 – Главное окно «Mouse auto clicker»

На рисунке 29 представлено главное окно программы «Mouse auto clicker». Окно компактное и позволяет записать нужные действия пользователя, а так же их воспроизвести с помощью двух кнопок. Так же имеется меню. Во вкладке «File» можно создать новый скрипт с помощью пункта «New», открыть уже созданный скрипт, пункт «Open» или сохранить записанный скрипт для автокликера с помощью пункта «Save». Пункт «Exit» завершает работу программы.

Вкладка «Options» позволяет регулировать скорость воспроизведения записанного скрипта с помощью пункта «Playback speed», что позволяет как ускорить, так и замедлить воспроизведение. Пункт «Recording options» позволяет задавать информацию, которую будет записывать скрипт, такую как нажатие кнопок на клавиатуре, движение мыши, клик мышкой, задержка между действиями.

Так же можно настраивать комбинации клавиш для управления программой с помощью пункта меню «Hotkey». Пункт «Startup» нужен для того, чтобы программа запускалась при загрузке операционной системы. Пункт «Other» предназначен для установки действий, после выполнения записанного скрипта.

Раздел «Help» предназначен для получения справки пользователем при работе с программой. Данная программа проста в использовании и помогает пользователю автоматизировать работу и упростить выполнение монотонных задач. Позволяет создавать сценарии часто повторяющихся действий.

## **4 Программное обеспечение естественно-языковых интерфейсов**

### **4.1 Разработка систем «Генератор учебных заданий» и «Генератор порождаемых шаблонов»**

На основе алгоритмов порождаемых шаблонов и многомерных баз данных выполнена программная реализация систем «Генератор учебных заданий» и «Генератор порождаемых шаблонов». Эти системы предназначены для:

1. автоматической генерации учебных материалов к тексту на английском языке;
2. автоматического построения шаблонов для последующей генерации различных тестовых заданий к тексту для учащихся.

Разработанные системы содержат алгоритмическое и программное ядро для построения естественно-языковых интерфейсов. Системы прошли апробацию на задачах обучения английского языка. На английском языке были сгенерированы учебные задания, для последующей проверки знаний учащихся.

Данный комплекс программных систем реализует следующую схему генерации учебных материалов.

1. Автоматическая генерация шаблонов для последующей генерации учебных заданий.
  - 1.1. Выбор нужного шаблона.
  - 1.2. Копирование выбранного шаблона в программу «Генератор учебных заданий».
2. Создание шаблонов на основе многомерной базы данных для генерации учебных заданий на основе порождаемых шаблонов.
  - 2.1. Производится обход дерева генерации.
  - 2.2. Выбирается нужный узел дерева генерации осмысленных шаблонов.
  - 2.3. Данный шаблон записывается в группу шаблонов.
3. Автоматическая генерация учебных материалов.

### 3.1. Цикл:

3.1.1. Выбор нужного шаблона.

3.1.2. Ввод текста, на основе которого нужно сгенерировать задания.

3.1.3. Автоматическая генерация учебных материалов.

3.2. Сохранение полученной информации в виде файла формата rtf.

3.3. Использование полученных вариантов заданий, например, для проверки знаний учащихся на уроке.

Данная алгоритмическая схема может служить основой для генерации учебных многовариантных заданий и большей частью реализована в программах «Генератор учебных заданий» и «Генератор порождаемых шаблонов».

## **4.2 Назначение системы «Генератор учебных заданий»**

На основе изложенных выше теоретических принципов разработана электронная версия генерации учебных заданий к тексту на английском языке, позволяющая создавать учебные задания для проверки знаний учащихся. Разработка системы имело целью решение двух основных практических задач:

1. Увеличение объемов генерируемых материалов, а так же снижение временных затрат преподавателя.

2. Генерация в короткое время различных индивидуальных учебных заданий для учащихся, с возможностью выбора шаблонов для генерации.

Подобная унификация и классификация учебных заданий, поможет повысить эффективность проверки знаний учащихся.

Отличительной особенностью базы данных в данной программе является наличие определенного количества размерностей. Группы слов в этой многомерной базе данных представлены файлами, названия которых содержат в себе вектора признаков многомерной классификации. Таким образом, математическая модель многомерного пространства получает простую

программную реализацию.

Отличительной особенностью порождаемых шаблонов в данной программе является наличие особого синтаксиса:

*{the [Зэ] этот} + {not [нОт] не} + {quick [кВик] быстрый} + {writer [рАйтэ] писатель}  
+ {cook [кУ:к] готовить} + -(e)s + {the [Зэ] этот} + {dish [дИш] блюдо}  
{one [ВАН] один} + {surprisingly [сунрАйзинли] удивительно} + {sick [сИк] больной} +  
{major [мЕйджэ] майор} + {fry [фрАй] жарить} + -(e)s + {the [Зэ] этот} + {veal [вИ:л]  
телятина}*

### 4.3 Функции системы «Генератор учебных заданий»

Генератор учебных заданий может обеспечить генерацию, просмотр, тренировку, а также печать полученных материалов в файл. Все эти функции доступны в главном окне программы.

Пользователь может проверить свои знания сразу в программе, используя закладку «генератор заданий». Во вкладке «генератор заданий» будут доступны сгенерированные задания, на основе загруженного текста.

Чтобы распечатать материалы к уроку необходимо их сохранить в файл, используя кнопку «сохранить задания» во вкладке «генератор заданий».

Чтобы найти нужное слово и просмотреть его транскрипцию и перевод, нужно воспользоваться закладкой «словарь».

Таблица 4.1 – Функции генератора учебных заданий

	Название операции	Назначение операции
1	Загрузка текста в программу	Для генерации заданий и дальнейшего их решения
2	Анализ текста	Анализ текста необходим для того, чтобы узнать количество предложений в тексте, количество слов в тексте и количество слов из словаря



Продолжение таблицы 4.1

	Название операции	Назначение операции
3	Поиск слова в словаре по английскому написанию	Для нахождения перевода слова, если известно его написание на английском языке
4	Поиск слова в словаре по русскому написанию	Для нахождения слова в словаре, если известно его написание на русском
5	Сохранение созданных материалов в файл	Для сохранения материалов на диск, а так же дальнейшей распечатки материалов

Окончание таблицы 4.1

#### 4.4 Состав и структура системы «Генератор учебных заданий»

Программа «Генератор учебных заданий» содержит следующие компоненты:

- многомерную базу данных
- шаблоны для создания материалов
- словарь

Рассмотрим далее приведенные компоненты программы подробнее.

Группа Шаблонов представляют собой файл, из которых генератор заданий осуществляет случайный выбор и выдает именно те задания, которые связаны с грамматической и лексической темой рассматриваемые на уроке. Предусмотрено постредактирование данных шаблонов.

В таблице 4.2 приведены основные компоненты программы «Генератор заданий», в первой колонке указан раздел словаря, во второй колонке указывается, на какой части главной формы программы располагается соответствующий раздел, а в третьей колонке указывается его функция и назначение.

Таблица 4.2 – Основные составляющие компоненты программы «Генератор учебных заданий»

<b>Название раздела</b>	<b>Местоположение раздела</b>	<b>Функция</b>
Файл	Главное меню	Загрузка текста в программу, а так же анализ текста
Настройки	Главное меню	Настройка шрифта, цвета фона а так же панели инструментов
Помощь	Главное меню	Информация о программе и как ей пользоваться
Открыть текст	Главное меню/Файл	Загрузка необходимого текста в программу
Начать анализ	Главное меню/Файл	Анализ текста на количество предложений в тексте, количество слов в тексте, а так же количество слов из словаря
Выход	Главное меню/Файл	Выход из программы. Завершение работы программы
Шрифт	Главное меню/Настройки	Изменение шрифта текста
Цвет фона	Главное меню/Настройки	Изменение фонового цвета в программе
Панель инструментов	Главное меню/Настройки	Включение и отключение панели инструментов
Отсчет времени	Главное меню/Настройки	Включение функции ограничения времени на решение заданий
Справка	Главное меню/Помощь	Описание функций программы
О программе	Главное меню/Помощь	Информация о программе и ее авторах

Продолжение таблицы 4.2

Название раздела	Местоположение раздела	Функция
Исходный текст	Закладка сверху	Окно в котором отображается загруженный текст
Генератор заданий	Закладка сверху	Окно для генерации заданий, а так же их решения
Словарь	Закладка сверху	Окно для поиска нужного слова в словаре

Окончание таблицы 4.2

#### 4.5 Входные и выходные данные

Входными данными программы «Генератор учебных заданий» является следующая информация

- а) информация из многомерных баз данных;
- б) информация, вводимая пользователем;
- с) данные представленные в виде таблиц, рисунков и т. д.

Выходными данными программы «Электронный Словарь» является следующая информация

- а) распечатываемые задания к уроку;
- б) наглядная информация на экране дисплея.

Далее приводятся две таблицы, поясняющие структуру входных и выходных данных описываемой программы.

Таблица 4.3 – Входные данные программы «Генератор учебных заданий»

Входные данные	Как и где вводятся	Куда записываются
Файловая эмуляция многомерной базы данных	При запуске программы	В память программы
Генерируемые шаблоны	При запуске программы	В память программы
Текст на английском языке	Вводится пользователем, через опцию меню Open file	В память программы

В таблице 4.3 приведены входные данные программы «Генератор учебных заданий». Выходная информация программы представлена в виде следующей таблицы.

Таблица 4.4 – Выходные данные программы «Генератор учебных заданий»

Выходные данные	Как и где организуются	Куда записываются
Сгенерированные задания в формате .rtf	Образуются в результате применения алгоритма генерации учебных заданий	Записываются в файл в папку «Results» и выводятся на экран
Повторение слов по словарю	На закладке «словарь»	Через интерфейс программной системы

В таблице 4.4 приведены выходные данные программы «Генератор учебных заданий». Эти данные содержат информацию, используемую для изучения иностранного языка и для генерации учебных материалов.

#### 4.6 Функционирование системы «Генератор учебных заданий»

Операции, выполняемые при помощи системы «Генератор учебных заданий», это ряд следующих функций:

1. Запуск программы.
2. Загрузка текста.
3. Создание учебных материалов.
4. Сохранение материалов в файл.
5. Поиск слова в словаре.

Для генерации заданий в программе используется метод гибридизации многомерных баз данных и порождаемых шаблонов. В соответствии с этим методом, преподаватель открывает программу, загружает в нее нужный текст и генерирует задания. При помощи порождаемых шаблонов, генерируемые задания не повторяются, что помогает сделать варианты заданий уникальными

и повысить эффективность проверки знаний учащихся.

Примеры работы рассматриваемой программы приведены в виде следующей таблицы

Таблица 4.5 – Задача – генерация учебных материалов при помощи программы «Генератор учебных заданий»

<b>Операция</b>	<b>Путь задания начала операции</b>	<b>Назначение или результат операции</b>
Запуск программы	Файл TaskGenerator.exe	Загрузка программы и баз данных
Открытие текста	На закладке «файл» открыть путь к нужному файлу	Для последующей генерации заданий
Анализ текста	Во вкладке «исходный текст»	Для анализа количества слов в тексте, количества предложений в тексте, количества слов из словаря
Генерация заданий	Во вкладке «генератор заданий»	Создание учебных материалов на основе загруженного текста
Нахождение слова в словаре по его написанию	Закладка «словарь», текстовое поле для ввода слова и таблица для выбора нужного значения слова	Нахождение слова в словаре по его написанию
Сохранить задания в файл	Во вкладке «генератор заданий», кнопка «сохранить задания»	Сохранение сгенерированных учебных заданий в файл

В таблице 4.5 задается последовательность действий, необходимых для генерации учебных материалов при помощи программы «Генератор учебных заданий».

#### **4.7 Программная система «Генератор порождаемых шаблонов»**

#### **4.7.1 Назначение программы «Генератор порождаемых шаблонов»**

Генератор порождаемых шаблонов является средством генерации шаблонов, используемых в дальнейшем для генерации учебных заданий на основе текста.

Программа «Генератор порождаемых шаблонов» имеет два основных режима работы:

1) Генерация компьютером случайных предложений на основе выбранных подстановочных шаблонов. На основе последовательного выбора таких конструкций можно получать семантически связные тексты или предложения с достаточно сложной структурой.

2) Выбор пользователем необходимого шаблона из дерева шаблонов и копирование шаблонов в генератор заданий для составления учебных заданий к тексту на незнакомом пользователю языке.

#### **4.7.2 Функции программы «Генератор порождаемых шаблонов»**

Функции программы «Генератор порождаемых шаблонов» приведены далее в виде следующей таблицы.

Таблица 4.6 – Функции системы «Генератор порождаемых шаблонов»

	<b>Название операции</b>	<b>Путь для задания соответствующей команды</b>	<b>Назначение операции</b>
1	Выбор шаблона	Главное окно	Выбор нужного шаблона из дерева шаблонов для последующей генерации заданий на его основе

Продолжение таблицы 4.6

	Название операции	Путь для задания соответствующей команды	Назначение операции
2	Подстановка слов в шаблон	Главное окно, правой кнопкой мыши по выбранному шаблону, из выпадающего списка пункт «open all words»	Возможность узнать перевод непонятных слов или их произношение
3	Показать предложение в отдельном окне	Главное окно, правой кнопкой мыши по выбранному шаблону, из выпадающего списка пункт «show»	Для наглядного вывода в отдельное окно сгенерированного шаблона

Окончание таблицы 4.6

В таблице 4.6 приводятся основные функции программы «Генератор порождаемых шаблонов». После названия команды указывается, где и как ее можно задать, после чего описывается назначение предлагаемой операции.

#### 4.7.3 Структура программы «Генератор порождаемых шаблонов»

Структура программы «Генератор порождаемых шаблонов» приводится далее в виде следующей таблицы.

В таблице 4.7 приводятся основные структурные элементы программы «Генератор порождаемых шаблонов», например, поле с деревом шаблонов и вывода конечной сгенерированной строки служат для отображения генерируемых шаблонов.

Таблица 4.7 – Структура программы «Генератор порождаемых шаблонов»

	<b>Функциональный элемент</b>	<b>Функция элемента</b>	<b>Изменяемые параметры</b>
1	Поле с деревом шаблонов	Отображение дерева шаблонов	Генерируемый шаблон
2	Поле для вывода сгенерированного шаблона	Отображение конечной сгенерированной строки из дерева шаблонов	Конечная строка сгенерированного шаблона

#### **4.7.4 Входные и выходные данные программы «Генератор порождаемых шаблонов»**

В программу «Генератор порождаемых шаблонов» загружаются следующие данные:

1. База данных ссылок на многомерную базу данных,
2. Многомерная база данных.

Выходными данными программы является следующая информация:

1. Сгенерированный шаблон с осмысленными английскими предложениями по какой-либо теме,
2. Осмысленные предложения.

#### **4.7.5 Описание работы программы «Генератор порождаемых шаблонов»**

Далее опишем пошаговое выполнение двух основных задач, для экспериментального решения которых предназначена программа «Генератор порождаемых шаблонов».



Таблица 4.8 – Задача: генерация шаблонов для последующей генерации на их основе учебных заданий при помощи программы «Генератор порождаемых шаблонов»

Операция	Путь задания начала операции	Назначение или результат операции
Выбрать шаблон для порождения предложений	Дерево шаблонов	Выбор подстановочного шаблона, на основе которого генерируются предложения
Генерация случайного предложения на английском языке	Двойное нажатие по выбранному шаблону	Генерация предложения на английском языке
Просмотр перевода и произношения	Кнопка «All words» выпадающего меню, при нажатии правой кнопки мыши	Проверка правильности понимания предложения
Просмотр сгенерированного шаблона	Кнопка «Show» выпадающего меню, при нажатии правой кнопки мыши	Возможность просмотра в отдельном окне сгенерированного предложения

В таблице 4.8 задается последовательность действий, необходимая для генерирования шаблонов при помощи программы «Генератор порождаемых шаблонов». Пользователю следует выбрать один шаблон из дерева шаблонов, нажать кнопку подстановки слов и проверить сгенерированное предложение.

**Вывод:** была разработана программа «Генератор учебных заданий», предназначенная для электронного создания учебных материалов. Программа «Генератор учебных заданий» может использоваться для проверки знаний учащихся и генерирования учебных материалов для урока.

Так же была разработана программа «Генератор порождаемых шаблонов». Сгенерированные шаблоны в данной программе, в дальнейшем используются в программе «Генератор учебных заданий», для генерации многовариантных учебных материалов.

## 5 Описание разработанных программ

По вышеизложенным принципам реализовано два программных продукта – «Генератор порождаемых шаблонов» и «Генератор учебных заданий».

### 5.1 Визуальное отображение программы «Генератор порождаемых шаблонов».

На рисунке 30 представлено главное окно программы «Генератор порождаемых шаблонов». Программа состоит из двух окон. Во втором окне выводится дерево генерации шаблонов. В данном окне выбирается нужный шаблон и генерируется конечная строка. Во втором окне можно вывести конечную строку для наглядности. Для этого нужно нажать право кнопкой мыши по конечной строке и выбрать из выпадающего меню пункт «show». Данное действие представлено на рисунке 31.

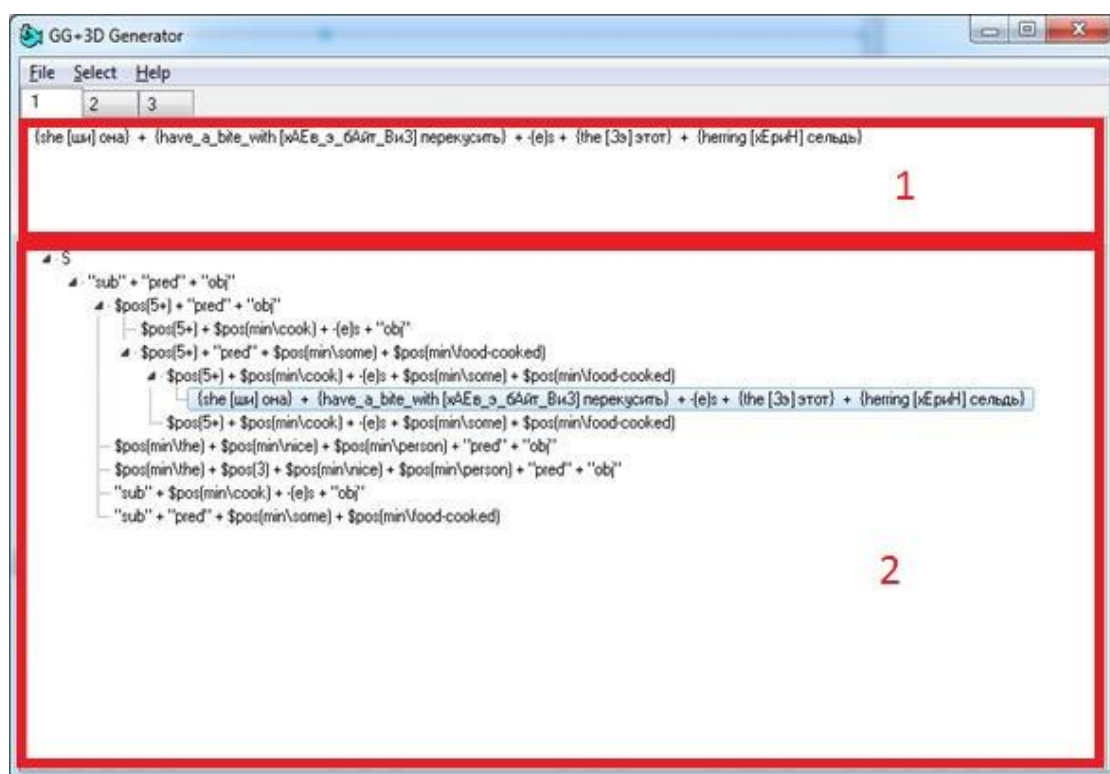


Рисунок 30 – Главное окно программы «Генератор шаблонов»

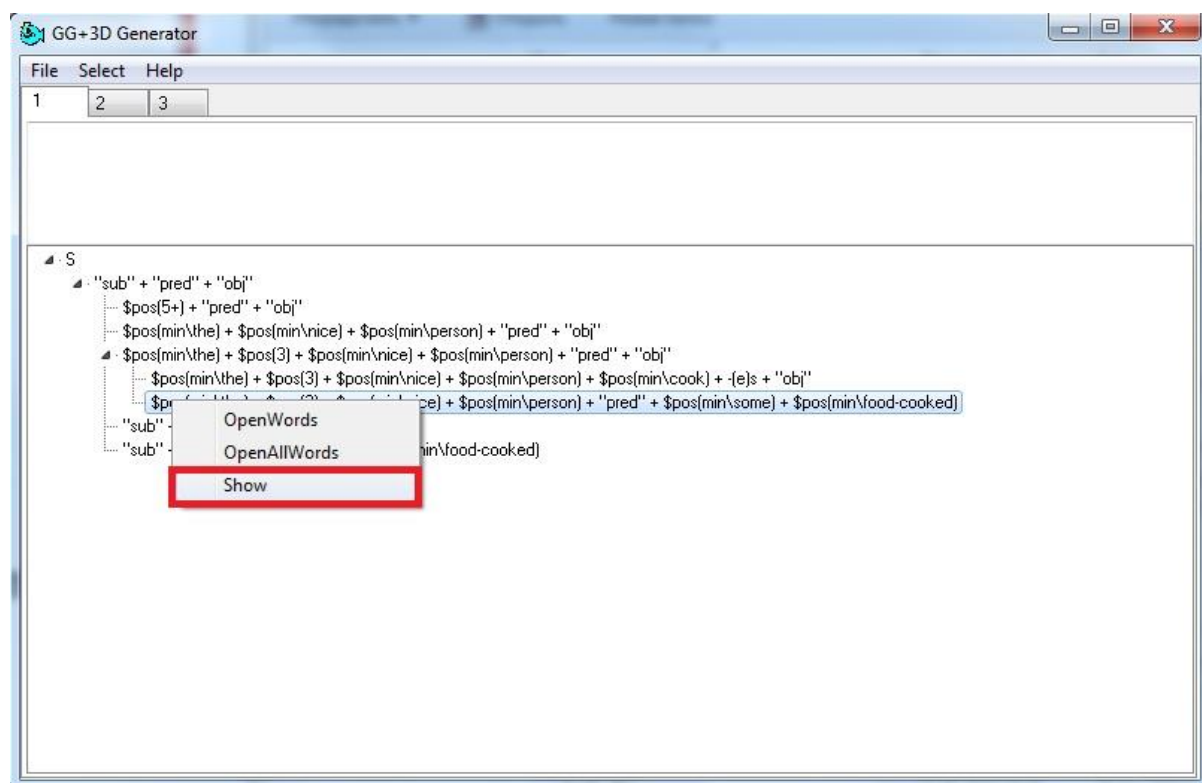


Рисунок 31 – Вывод строки в отдельном окне

## 5.2 Визуальное отображение программы «Генератор учебных заданий»

«Генератор учебных заданий» – это программный продукт, который обладает дружелюбным интерфейсом и ориентирован на пользователей, обладающих различным уровнем работы с компьютером. Данная программа получает в качестве входного параметра строку произвольного текста на английском языке с целью проанализировать его и обеспечить генерацию учебных заданий.

На рисунке 32 представлено окно программы «Генератор учебных заданий». В данную программу загружается произвольный текст в окно «исходный текст» на английском языке с целью проанализировать его и сгенерировать учебные задания. Чтобы загрузить текст, нужно в панели меню выбрать пункт «Файл — Открыть текст». Затем нужно нажать кнопку «Начать анализ», для того, чтобы проанализировать сколько в тексте предложений и

слов, а так же узнать количество слов из словаря. После того, как текст будет проанализирован, можно переходить во вкладку «Генератор заданий» и генерировать учебные задания.

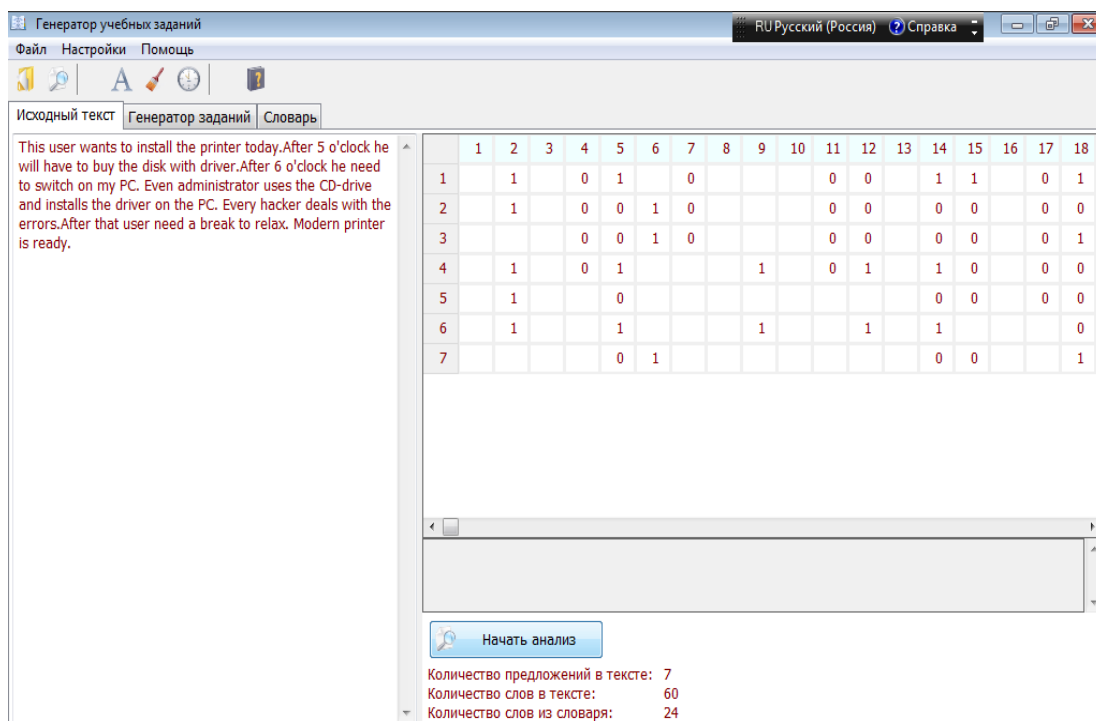


Рисунок 32 – Главное окно программы с загруженным текстом и его визуализацией

На основе вариантов ответов, сгенерированных с помощью программы можно составлять задания типа multiple choice (выбор одного варианта ответа из многих), multiple select (выбор нескольких правильных ответов из многих), rearrangement (задания на упорядочение), matching (задания на перекрестный выбор), multiple matching (задания на множественное соответствие) и др.

Задания 1 типа: нужно заполнить пропуск в предложении. Дано 4 варианта и из этих вариантов нужно выбрать подходящее слово по смыслу и заполнить пропуск в предложении.

Задания типа 2: нужно заполнить пропуск гласной буквы в слове. Дается пять вариантов и из них нужно выбрать одну гласную и заполнить ей пропуск в слове.

Задания типа 3: нужно перевести слово на русский язык и выбрать правильный вариант из четырех предложенных вариантов.

Задания типа 4: нужно напечатать перевод слова строчными буквами на русский язык.

Задания типа 5: Печать перевода слова на английском языке. Нужно напечатать строчными буквами перевод на английском языке слова написанного на русском языке.

Примеры заданий представлены на рисунке 33.

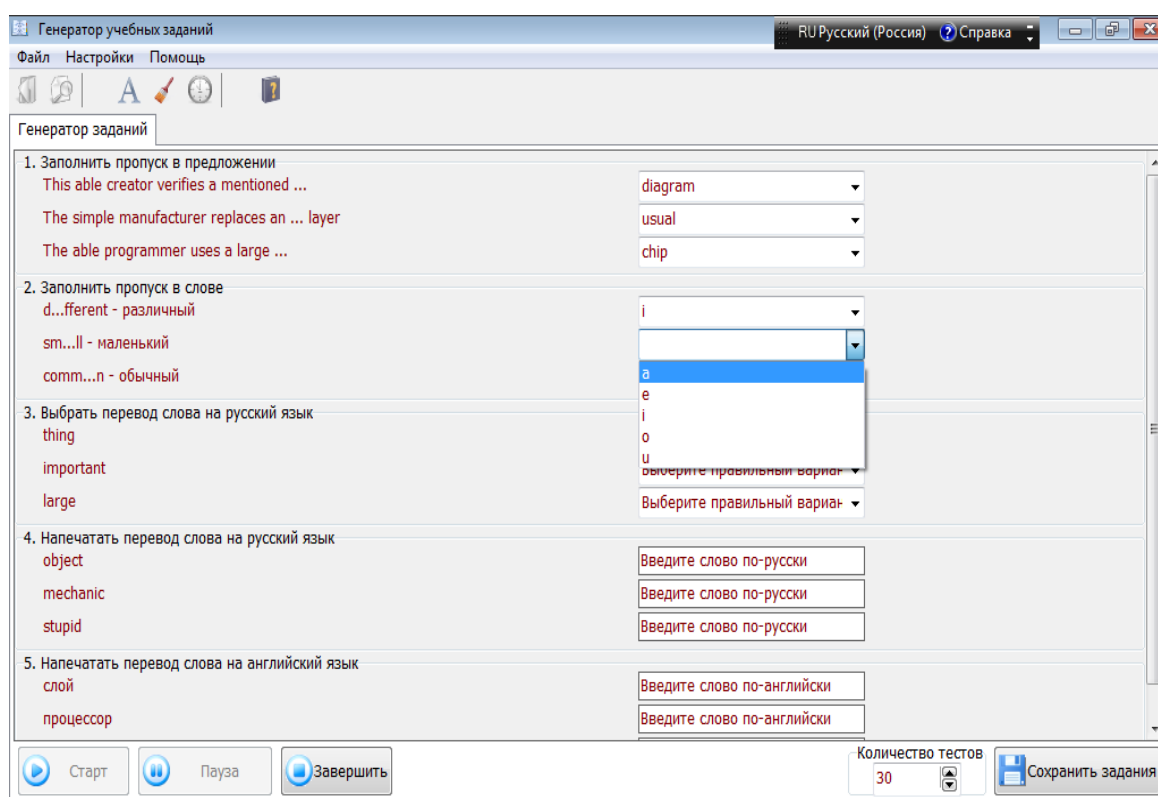


Рисунок 33 – Программа после выполнения проверки заданий

В нижней части вкладки находятся кнопки управления:

«Старт» – генерация новых заданий с выводом на экран;

«Пауза» – приостановка выполнения заданий;

«Завершить» – окончание выполнения заданий и проверка правильности ответов.

У пользователя есть возможность решить сгенерированные задания в программе. Для этого нужно нажать кнопку «Старт».

Так же можно сохранить задания в файл формата \*.rtf. Для того что бы сохранить задания нужно выбрать количество заданий и нажать кнопку в нижнем правом углу «Сохранить задания». Располагается сохраненный файл в директории TaskGenerator/Results/. Пример сохраненного файла представлен на рисунке 34.

Тестовое задание 013 (от 2015-09-16)					
<b>1. Заполнить пропуск в предложении</b>					
The stupid programmer verifies a following ...	<input type="checkbox"/> application	<input type="checkbox"/> chicken			
	<input type="checkbox"/> coco	<input type="checkbox"/> turkey			
That ... creator uses a huge application	<input type="checkbox"/> solid	<input type="checkbox"/> talented			
	<input type="checkbox"/> wrong	<input type="checkbox"/> analogue			
The able mechanic connects an ... thing	<input type="checkbox"/> sour	<input type="checkbox"/> sweet			
	<input type="checkbox"/> analogue	<input type="checkbox"/> salty			
<b>2. Заполнить пропуск в слове</b>					
comm__n - обычный	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
r__gister - регистр	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
ord__r - заказывать	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
<b>3. Выбрать перевод слова на русский язык</b>					
compiler	<input type="checkbox"/> компилятор	<input type="checkbox"/> наладонник			
	<input type="checkbox"/> катод	<input type="checkbox"/> диод			
huge	<input type="checkbox"/> огромный	<input type="checkbox"/> цвета морской волны			
	<input type="checkbox"/> ярко-красный	<input type="checkbox"/> большой			
complex	<input type="checkbox"/> прозрачный	<input type="checkbox"/> обычный			
	<input type="checkbox"/> сложный	<input type="checkbox"/> хаки			
<b>4. Напишите перевод слова на русский язык</b>					
mechanic					
element					
stupid					
<b>5. Напишите перевод слова на английский язык</b>					
создатель					
обычный					
программист					

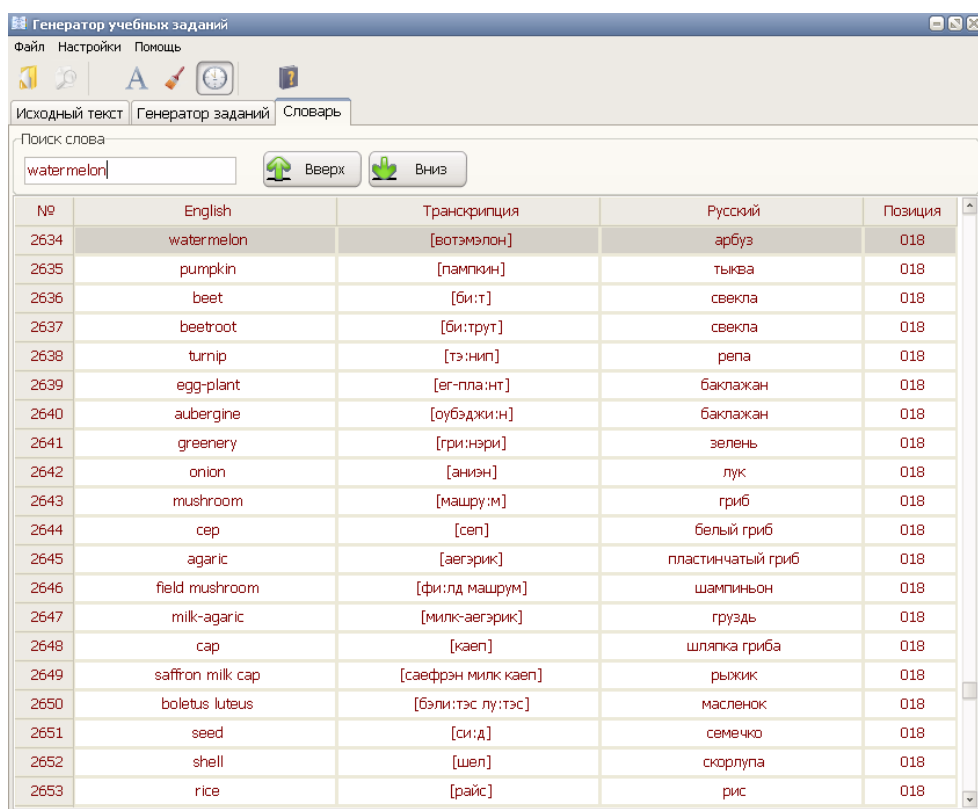
Рисунок 34 – Сохранение сгенерированных заданий

В программе, в помощь пользователю, имеется вкладка «Словарь». В словаре имеется возможность поиска нужных слов, как на английском, так и на русском языках.

Файлы словаря располагаются в папке TaskGenerator\Base\.

Во вкладке содержится таблица, в которой содержится 5 колонок. В первой колонке содержится порядковый № слова, во второй колонке содержится слово на английском языке, в третьей транскрипция, в четвертой перевод слова на русском языке и в пятой колонке содержится позиция слова в предложении.

Чтобы найти нужное слово в словаре, нужно ввести слово хоть на английском, хоть на русском языке в поле «Поиск слова» и затем с помощью кнопок «Вверх» и «Вниз» искать слово в нужном направлении. Вид словаря изображен на рисунке 35.



№	English	Транскрипция	Русский	Позиция
2634	watermelon	[вотэмэлон]	арбуз	018
2635	pumpkin	[пампкин]	тыква	018
2636	beet	[би:т]	свекла	018
2637	beetroot	[би:трут]	свекла	018
2638	turnip	[тэ:нип]	репа	018
2639	egg-plant	[ег-плант]	баклажан	018
2640	aubergine	[оубэджин]	баклажан	018
2641	greenery	[гри:нэри]	зелень	018
2642	onion	[аниэн]	лук	018
2643	mushroom	[машру:м]	гриб	018
2644	cep	[сеп]	белый гриб	018
2645	agaric	[аегэрик]	пластинчатый гриб	018
2646	field mushroom	[фи:лд машрум]	шампиньон	018
2647	milk-agaric	[милк-аегэрик]	груздь	018
2648	cap	[каеп]	шляпка гриба	018
2649	saffron milk cap	[саефрэн милк каеп]	рыжик	018
2650	boletus luteus	[бэли:тэс лу:тэс]	масленок	018
2651	seed	[си:д]	семечко	018
2652	shell	[шел]	скорлупа	018
2653	rice	[райс]	рис	018

Рисунок 35 – Англо-русский словарь

## **ЗАКЛЮЧЕНИЕ**

Была разработана программа «Генератор учебных заданий», а так же модуль, который позволяет осуществить гибридизацию многомерных баз данных и порождающих грамматик. «Генератор учебных заданий», позволит автоматизировать работу преподавателей, а так же ускорить процесс разработки тестов.

Так же была разработана программа «Генератор порождаемых шаблонов», которая позволяет создавать различные шаблоны для генерации заданий.

Получены следующие результаты: были разработаны правила порождающей грамматики, многомерные базы данных на основе бесплатных источников, разработаны модули, которые позволяют осуществить гибридизацию многомерных баз данных и порождающих грамматик, а так же программа генерации шаблонов, позволяющая порождать учебные задания по английскому языку на основе входного текста.

Была предложена модель гибридизации многомерных баз данных и порождаемых шаблонов. Данная модель помогает сделать задания уникальными, а так же разнообразить их содержание.

На основе предложенной модели был разработан генератор заданий, который позволяет автоматически генерировать учебные материалы, при этом данная программа помогает экономить время преподавателя, а благодаря множеству различных вариантов заданий, предотвратить списывание и повысить качество обучения.

Данные продукты могут быть внедрены в будущем в высших учебных заведениях, школах и других организациях.



## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Borko, H. Abstracting Concepts and Methods / H. Borko, C.L. Bernier // – New York: Academic Press, 1975. – P. 140.
2. Comrie, B. Language Universals and linguistic typology / B. Comrie // University Of Chicago Press, 2nd edition, 1989. – P.230-239.
3. Cornelius, T. Leondes Intelligent systems: technology and applications. / T. Cornelius // – CRC Press, 2002. – P. 127-128.
4. Everett, D. R. Computer-Mediated Communication as a Teaching Tool: a Case Study / D. R. Everett, T.C. Ahern // Journal of Research on Computing in Education, Vol.26, Iss.3, 1994. – P. 336-337.
5. Francis, W. N. Manual of Information to accompany A Standard Corpus of Present-Day Edited American English, for use with Digital Computers / W. N. Francis, H. Kucera // – Brown University Providence, Rhode Island Department of Linguistics Brown University, 1979 – P. 18-21.
6. Hahn, U. Knowledge-Based Text Summarization: Salience and Generalization Operators for Knowledge-Based Abstraction / U. Hahn, U. Reimer // – USA, Massachussets, Cambridge, MIT Press, 1999. – P. 215-232.
7. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. / J.J. Hopfield // – National Academy Science, 1982. – P. 254-258.
8. Hutchins, J. Summarization: Some Problems and Methods / J. Hutchins, K. Jones // Informatics: Meaning the Frontier of Informatics. – London, Aslib, 1987. – P. 151-173.
9. Ingraham, B. Language Training for Various Purposes in Several Languages on a Common Hypermedia Framework / B. Ingraham, T. Chanier, C. Emery // Computer & Education, Vol.23, Iss. 1,2, 1994. – P. 107-115.
10. Jones, G. Computer Simulations in Language Teaching – the Kingdom Experiment / G. Jones // System, Iss.17, 1986. – P. 35-47.

11. Kupiec, J. A Trainable Document Summarizer / J. Kupiec, J. Pedersen, F. Chen // 18th International ACM SIGIR Conference: Research and Development in Information Retrieval. – New York, ACM Press, 1995. – P. 68-73.
12. Levy, M. Computer Assisted Language Learning: Context and Conceptualization / M. Levi // – Oxford: Clarendon Press, 1997. – P. 130-133.
13. Mani, I. Text Summarization Evaluation / I. Mani // 9th European Conference, 2000. – P. 101-107.
14. McQuail, D. Models of communication / D. McQuail, S. Windahl // – New York: Oxford University Press, vol. 3, pp. 36-44, 1989.
15. Nagata, N. Computer vs. Wordbook Instruction in Second Language Acquisition / N. Nagata // CALICO Journal, Vol.14, №1, 1996. – P. 53-76.
16. Radev, D.R. Generating Natural Language Summaries from Multiple Online Sources / D.R. Radev, K.R. McKeown, // Computational Linguistics, Vol. 24, No. 3, 1998. – P. 469-500.
17. Salton, G. Automatic Text Structuring and Summarization / G. Salton // Information Processing & Management, Vol. 33, No. 2, 1997. – P. 193-207.
18. Shannon, C.E. A Mathematical Theory of Communication / C. E. Shannon // The Bell System Technical Journal, №27, 1948. – P. 379-423, 623-656.
19. Smith, S. Educational Uses of the PLATO Computer System / S. Smith, B. Sherwood // – Science, 1976. – P. 82-83.
20. Weaver, W. The mathematics of communication / W. Weaver // Scientific American, Vol.181, Iss.11-15, 1949. – P. 30-40.
21. Wyllys, R.E. Extracting and Abstracting by Computer Automated Language Processing / R.E. Wyllys, H. Borko, John Wiley // – New York, 1967. – P. 127-179.
22. Агапова, О.И. О трех поколениях компьютерных технологий обучения / О.И. Агапова, А.О. Кривошеев, А.С. Ушаков // Информатика и образование, №2, 1994. – С. 34-40.
23. Азимов, Э.Г. К типологии обучающих компьютерных программ по русскому языку как иностранному. Современные технические средства в

обучении русскому языку как иностранному / Э.Г. Азимов // – Москва, 1990. – С. 136-143.

24. Айвазян С.А. Прикладная статистика. Статистическое оценивание зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин // 2-е изд. – Москва: Финансы и статистика, 1991.- 512 с.

25. Айвазян, С.А. Прикладная статистика. Статистическое оценивание зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин // – Москва: Финансы и статистика, 1-е изд-е, 1985. – 484 с.

26. Александров, Г.Н. Программированное обучение и новые информационные технологии обучения / Г.Н. Александров // Информатика и образование, №5, 1993. – С.7 -19.

27. Атанов, Г.А. Структурирование понятий предметной области с помощью методов представления знаний. / Г.А. Атанов, И.Н. Пустынникова // Искусственный интеллект, №2, 1997. - С. 29-40.

28. Башмаков, А.И. Технология и инструментальные средства проектирования тренажерно-обучающих комплексов для профессиональной подготовки и повышения квалификации. Часть 1. / А.И. Башмаков // Информационные технологии, №6, 1999. – С. 40-45.

29. Башмаков, А.И. Технология и инструментальные средства проектирования тренажерно-обучающих комплексов для профессиональной подготовки и повышения квалификации. Часть 2. / А.И. Башмаков // Информационные технологии, №7, 1999. – С. 39-45.

30. Бирюков, Р.А. Некоторые методы автоматического реферирования / Р.А. Бирюков // Открытые Системы, № 12, 2000. – С. 30-35.

31. Бовтенко, М.А. Компьютерная лингводидактика: учеб.пособие / М.А. Бовтенко // – Новосибирск: Изд-во НГТУ, 2000. – С. 45-47.

32. Брусиловский, П.Л. Интеллектуальные обучающие системы / П.Л. Брусиловский // Информатика. Информационные технологии. Средства и системы, №2, 1990. – С. 3-22.

33. Власов, Е.А. Компьютеры в обучении языку: проблемы и решения / Е.А. Власов, Т.Ф. Юдина // – Москва, 1990. – 112 с.
34. Дейвисон, М. Многомерное шкалирование: методы наглядного представления данных / М. Дэйвисон // – Москва: Финансы и статистика, 1988. – 154 с.
35. Довгялло, А.М. Обучающие системы нового поколения / А.М. Довгялло, Е.Л. Ющенко // УсиМ, №1, 1988. – С. 83-86.
36. Домрачев, В.Г. О классификации компьютерных образовательных информационных технологий / В.Г. Домрачев, И.В. Ретинская // Информационные технологии, №2, 1996. – С. 10-14.
37. Жарков, И.В. Автоматизированные обучающие системы / И.В. Жарков, под ред. Гердт А.С. // – СПб., 1996. – С. 59-68.
38. Зиновьев, А.Ю. Визуализация многомерных данных: монография / А.Ю. Зиновьев // Красноярск: Изд-во КГТУ, 2000. – 168 с.
39. Кларлащук, В.И. Обучающие программы / В.И. Кларлащук // – Москва: Солон-Р, 2001. – 94 с.
40. Кручинин, В.В. Разработка компьютерных учебных программ / В.В. Кручинин // – Томск, 1998. – 120 с.
41. Лаутербах, Р.Н. Программное обеспечение процесс обучения. Перспективы / Р.Н. Лаутербах, К.К. Фрей // Вопросы образования, 1988, №3. – С. 70-79.
42. Лбов, Г.С. Методы обработки разнотипных экспериментальных данных / Г.С. Лбов // – Новосибирск: Наука, 1981. – 157 с.
43. Личаргин, Д. В. К проблеме генерации расширенной фонетической транскрипции на основе векторных порождающих грамматик. [Электронный ресурс] / Д. В Личаргин, В. В. Сотникова, А. А. Усова, В. В. Бутовченко // Современные проблемы науки и образования. – 2015. № 2. . – Режим доступа: <http://www.science-education.ru/ru/article/view?id=20933>.
44. Личаргин Д. В. Разработка веб-приложения по генерации учебных заданий к тексту на естественном языке на основе порождаемых шаблонов. / В.

В. Сотникова, А. А. Усова, И. В. Колбасина, Е. Д. Старовойт // Новосибирск: X Международная научно-практическая конференция: «Научные перспективы XXI века. Достижения и перспективы нового столетия». – 2015. – С. 57-60.

45. Личаргин Д.В. Разработка приложения по генерации учебных заданий к тексту на естественном языке на основе порождаемых шаблонов. / А. А. Усова, В. В. Сотникова, С. А. Липман, В. В. Бутовченко. // Современные проблемы науки и образования. – 2015. № 6. – Режим доступа: <http://www.science-education.ru/ru/article/view?id=22636>.

46. Маклюэн, М. Понимание медиа: внешние расширения человека / М. Маклюэн // – Москва, 2003. – С.24-29.

47. Машбиц, Е.И. Диалог в обучающей системе / Е.И. Машбиц // – Киев, 1989. – 88 с.

48. Назаров, М.М. Массовая коммуникация и общество. Введение в теорию и исследования / М.М. Назаров // – Москва, 2004. – 235 с.

49. Науменко, Т.В. Социология массовой коммуникации / Т.В. Науменко // – Москва, 2005. – 167 с.

50. Ножов, И.М. Синтаксический анализ / И.М. Ножов // Компьютерра, № 21 (446), 2002 – С.42-49.

51. Ножов, И.М. Морфологическая и синтаксическая обработка текста (модели и программы): интернет публикация / И.М. Ножов // – Москва, 2003. – 22 с.

52. Околелов, Г.Н. Программированное обучение и новые информационные технологии обучения / Г.Н. Околелов // Информатика и образование, №5, 1993. – С. 7-19.

53. Полинченко, Д.Ю. Естественный язык как лингвокультурный семиотический концепт: автореф. дис... канд. филол. наук / Д.Ю. Полинченко // – Волгоград, 2004. – 22 с.

54. Растрингин, Л.А. Адаптация сложных систем. Методы и приложения / Л.А. Растрингин // – Рига: Зинатне, 1981, – 124 с.

55. Савельев, А.Я. Подготовка информации для автоматизированных обучающих систем / А.Я. Савельев, В.А. Новиков, Ю.И. Лобанов // – Москва: Высшая школа, 1986. – 243 с.
56. Соловов, А.В. Проектирование компьютерных систем учебного назначения: учебное пособие / А.В. Соловов // – Самара: СГАУ, 1995.
57. Сотникова В. В. Векторные порождающие грамматики и их применение для генерации учебных материалов по фонетике языка. / В. В. Сотникова, А. А. Усова, Д. В. Личаргин //Красноярск: Международная конференция студентов, аспирантов и молодых ученых «Перспектив-2015», посвященная 70-летию Великой Победы. – 2015. – С. 0-92.
58. Терехина, А.Ю. Анализ данных методами многомерного шкалирования / А.Ю. Терехина // – Москва: Наука, 1986. – 168 с.
59. Терехов, С.А. Нейросетевые информационные модели сложных инженерных систем. Нейроинформатика / С.А. Терехов // – Новосибирск. Наука. 1998. – С.101-136.
60. Усова А. А. Разработка веб-приложения по генерации учебных заданий к тексту на естественном языке на основе порождаемых шаблонов. / А. А. Усова, В. В. Сотникова, Д. В. Личаргин // Красноярск: Международная конференция студентов, аспирантов и молодых ученых «Перспектив-2015», посвященная 70-летию Великой Победы. – 2015. – С. 7-10.

## ПРИЛОЖЕНИЕ А

Задания сгенерированные программой «Генератор учебных заданий»

### Тестовое задание 001

#### 1. Заполнить пропуск в предложении

This simple user purchases a sophisticated ...	<input type="checkbox"/> coconut	<input type="checkbox"/> dish
	<input type="checkbox"/> sturgeon	<input type="checkbox"/> text
This responsible chemist ... a neural library	<input type="checkbox"/> freezes	<input type="checkbox"/> grates
	<input type="checkbox"/> replaces	<input type="checkbox"/> shakes
The clever student sets up a digital ...	<input type="checkbox"/> sausage	<input type="checkbox"/> caviare
	<input type="checkbox"/> monitor	<input type="checkbox"/> agaric

#### 2. Заполнить пропуск в слове

c___mplex - сложный	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
sc___entist - ученый	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
engine___r - инженер	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u

#### 3. Выбрать перевод слова на русский язык

computer	<input type="checkbox"/> анод	<input type="checkbox"/> струйный принтер
	<input type="checkbox"/> компьютер	<input type="checkbox"/> photoshop
replace	<input type="checkbox"/> проверять	<input type="checkbox"/> уничтожать
	<input type="checkbox"/> делать	<input type="checkbox"/> заменить
telephone	<input type="checkbox"/> передний план	<input type="checkbox"/> регистр
	<input type="checkbox"/> цветной телевизор	<input type="checkbox"/> телефон

#### 4. Напишите перевод слова на русский язык

front	_____
various	_____
compiler	_____

#### 5. Напишите перевод слова на английский язык

объект	_____
правый	_____
женщина	_____

## Тестовое задание 002

### 1. Заполнить пропуск в предложении

The simple student prepares a ... processor	<input type="checkbox"/> sweet	<input type="checkbox"/> tasteless
	<input type="checkbox"/> flavoured	<input type="checkbox"/> sophisticated
The clever british needs a ... processor	<input type="checkbox"/> balmy	<input type="checkbox"/> tinned
	<input type="checkbox"/> digital	<input type="checkbox"/> warm
This clever ... holds a right diode	<input type="checkbox"/> discount	<input type="checkbox"/> duty
	<input type="checkbox"/> target	<input type="checkbox"/> mechanic

### 2. Заполнить пропуск в слове

c__rcuit - цепь	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
comput__r - компьютер	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u
displ__y - дисплей	<input type="checkbox"/> a	<input type="checkbox"/> e	<input type="checkbox"/> i	<input type="checkbox"/> o	<input type="checkbox"/> u

### 3. Выбрать перевод слова на русский язык

phone	<input type="checkbox"/> телефон	<input type="checkbox"/> дисплей
	<input type="checkbox"/> пейджер	<input type="checkbox"/> окно
control	<input type="checkbox"/> запускать	<input type="checkbox"/> проектировать
	<input type="checkbox"/> управлять	<input type="checkbox"/> соединять
connect	<input type="checkbox"/> разрушать	<input type="checkbox"/> соединять
	<input type="checkbox"/> подключать	<input type="checkbox"/> приносить

### 4. Напишите перевод слова на русский язык

simple	_____
desktop	_____
clever	_____

### 5. Напишите перевод слова на английский язык

правый	_____
цифровой	_____
передний	_____



## Тестовое задание 003

### 1. Заполнить пропуск в предложении

That responsible technician ... a special  
circuit

☐ controls

☐ eats

☐ pickles

☐ tins

This simple technician breaks a complex ...

☐ almond

☐ circuit

☐ second course

☐ beef

The ... manufacturer owns a blue processor

☐ wooden

☐ mini

☐ clever

☐ neural

### 2. Заполнить пропуск в слове

l\_\_\_rge - большой

☐ a

☐ e

☐ i

☐ o

☐ u

c\_\_\_mputer - компьютер

☐ a

☐ e

☐ i

☐ o

☐ u

displ\_\_\_y - дисплей

☐ a

☐ e

☐ i

☐ o

☐ u

### 3. Выбрать перевод слова на русский язык

american

☐ ребенок

☐ американец

☐ ритор

☐ биолог

simple

☐ помешанный

☐ умелый

☐ простой

☐ беззаботный

physicist

☐ новичок

☐ физик

☐ доктор

☐ библиотекарь

### 4. Напишите перевод слова на русский язык

special

programmer

electromagnet

### 5. Напишите перевод слова на английский язык

студент

умный

проект

## ПРИЛОЖЕНИЕ Б

Шаблоны сгенерированные программой «Генератор порождаемых шаблонов»

{I [ай] Я} + {cook [кУ:к] готовить} + -(e)s + {the [Зэ] этот} + {buckwheat [бАкВИ:т] гречка}

{we [ви] мы} + {shred [шрЕд] шинковать} + -(e)s + {the [Зэ] этот} + {vegetable [вЕджетэбэл] овощ}

{he [хи] он} + {fry [фрАй] жарить} + -(e)s + {the [Зэ] этот} + {pearl-barley [пЭ:л\_бА:ли] перловая\_крупa}

{you [йиу] ты,Вы} + {eat [И:т] есть} + -(e)s + {that [ЗАЕт] тот} + {cutlet [кАтлет] котлета}

{I [ай] Я} + {eat [И:т] есть} + -(e)s + {the [Зэ] этот} + {wheat [ВИ:т] пшеница}

{you [йиу] ты,Вы} + {shred [шрЕд] шинковать} + -(e)s + {the [Зэ] этот} + {food [фУ:д] пища}

{she [ши] она} + {congeal [конджИ:л] заморозить} + -(e)s + {the [Зэ] этот} + {sausage [сОсидж] сосиска}

{she [ши] она} + {roast [рОуст] тушить} + -(e)s + {this [ЗИс] этот} + {poultry [пАултри] дичь}

{you [йиу] ты,Вы} + {eat [И:т] есть} + -(e)s + {some [сАм] немного} + {cutlet [кАтлет] котлета}

{she [ши] она} + {cook [кУ:к] готовить} + -(e)s + {this [ЗИс] этот} + {first\_course [фЭ:ст\_кО:с] первое}

{I [ай] Я} + {cook [кУ:к] готовить} + -(e)s + {this [ЗИс] этот} + {dessert [дизЭ:т] десерт}

{we [ви] мы} + {have\_a\_bite\_with [хАЕв\_э\_бАйт\_ВиЗ] перекусить} + -(e)s + {the [Зэ] этот} + {meat [МИ:т] мясо}

{we [ви] мы} + {eat [И:т] есть} + -(e)s + {the [Зэ] этот} + {carrot [кАЕрэт] морковь}

{he [хи] он} + {taste [тейст] пробовать} + -(e)s + {the [Зэ] этот} + {dish [дИш] блюдо}

{he [хи] он} + {roast [рОуст] тушить} + -(e)s + {that [ЗАЕт] тот} + {сer [сЕп] белый\_гриб}

{he [хи] он} + {eat [И:т] есть} + -(e)s + {one [ВАН] один} + {garnish [гА:ниш] гарнир}

{she [ши] она} + {have [хАЕв] есть,пить} + -(e)s + {the [Зэ] этот} + {dessert [дизЭ:т] десерт}

{I [ай] Я} + {warm\_up [ВОм\_ап] разогревать} + -(e)s + {the [Зэ] этот} + {pork [пО:к] свинина}

{I [ай] Я} + {cut [кАт] резать} + -(e)s + {the [Зэ] этот} + {cutlet [кАтлет] котлета}

{we [ви] мы} + {have [хАЕв] есть,пить} + -(e)s + {the [Зэ] этот} + {egg-plant [Ег-плА:нт] баклажан}

{we [ви] мы} + {cook [кУ:к] готовить} + -(e)s + {that [ЗАЕт] тот} + {mushroom [мАшрУ:м] гриб}

{you [йиу] ты,Вы} + {eat [И:т] есть} + -(e)s + {this [ЗИс] этот} + {sausage [сОсидж] сосиска}

{I [ай] Я} + {eat [И:т] есть} + -(e)s + {the [Зэ] этот} + {pearl-barley [пЭ:л\_бА:ли] перловая\_крупа}

{a(n) [э(н)] некоторый} + {pedantic [пидАЕнтик] педантичный} + {lady [лЕйди] леди} + {fry [фрАй] жарить} + -(e)s + {the [Зэ] этот} + {beef [бИ:ф] говядина}

{ {none [нОу] никакой} + of + the [Зэ] этот} + {weak [ВИ:к] слабый} + {active\_worker [АЕктив\_ВЭ:кэ] активист} + {warm\_up [ВОм\_ап] разогревать} + -(e)s + {the [Зэ] этот} + {vegetable [вЕджетэбэл] овощ}

{ {a\_few [фйУ ☺ мало} + of + the [Зэ] этот} + {tall [тОл] высокий} + {girl [гЭ:л] девушка} + {refrigerate [рефриджэрЕйт] заморозить} + -(e)s + {the [Зэ] этот} + {beef [бИ:ф] говядина}

{that [ЗАЕт] тот} + {powerful [пАуэфул] влиятельный} + {sailor [сЕйлэ] моряк} + {satisfy\_hunger\_with [сАЕтисфай\_хАнгэ\_ВИЗ] утолять\_голод} + -(e)s + {this [ЗИс] этот} + {potatoes [потЕйтэз] картофель}

{one [ВАН] один} + {active [АЕктив] активный} + {chemist [кЕмист] химик} + {have [хАЕв] есть,пить} + -(e)s + {the [Зэ] этот} + {onion [Аниэн] лук}

{that [ЗАЕт] тот} + {strong [стрО:н] сильный} + {magician [мэджИшн] волшебник} + {fry [фрАй] жарить} + -(e)s + {this [ЗИс] этот} + {food [фУ:д] пища}

{ {some [сАм] немного} + of + the [Зэ] этот} + {disobedient [дисэБИдиэнт] непослушный} + {lady [лЕйди] леди} + {cut [кАт] резать} + -(e)s + {the [Зэ] этот} + {sausage [сОсидж] колбаса}

{the [Зэ] этот} + {clever [клЕвэ] умный} + {king [КИН] король} + {warm\_up [ВОм\_ап] разогреть} + -(e)s + {this [ЗИс] этот} + {macaroni [маекэрОуни] макароны}

{the [Зэ] этот} + {sensitive [сЕнситив] чувствительный} + {grown-up\_person [грОун-ап] взрослый} + {cool [кУ:л] охладить} + -(e)s + {that [ЗАЕт] тот} + {mushroom [мАшрУ:м] гриб}

{this [ЗИс] этот} + {handsome [хАЕндсэм] красивый} + {lieutenant [лэфтЕнаэт] лейтенант} + {cook [кУ:к] готовить} + -(e)s + {this [ЗИс] этот} + {sausage [сОсидж] колбаса}

{ {both [бОуС] оба} + of + the [Зэ] этот} + {leisurely [лЕжэли] неторопливый} + {pedestrian [педЕстриэн] прохожий} + {have [хАЕв] есть,пить} + -(e)s + {the [Зэ] этот} + {mutton [мАтн] баранина}